THERAPEUTICS AND DIAGNOSTICS FOR OCULAR ABNORMALITIES

Ins A1

BACKGROUND OF THE INVENTION

The normal anatomic relationship between the retina and the choroid is crucial to eye function. In many ocular diseases or abnormalities such as, for example, chorioretinal degenerations, retinal degenerations, macular degenerations and retinal detachment, the integrity of this relationship is compromised. One of the most common ocular disorders is macular degeneration (MD), a clinical term that is used to describe a variety of diseases that are all characterized by a progressive loss of central vision associated with abnormalities of Bruch's membrane and the retinal pigment epithelium (RPE) (See Figure 1). These disorders include very common conditions that affect older patients (age-related macular degeneration, or AMD), as well as rarer, earlier-onset dystrophies. Age-related macular degeneration (AMD) is the leading cause of legal blindness in the United States, Canada, and the United Kingdom. AMD is associated with progressive diminution of visual acuity in the central portion of the visual field, changes is color vision, and abnormal dark adaptation and sensitivity. These deficits are due to the degeneration and eventual loss of photoreceptor and RPE cells within the macula, a highly specialized relatively avascular region of the central retina containing a high density of cone photoreceptor cells and the cone-rich fovea. AMD is also clinically characterized, in part, by the presence of extracellular deposits, called "drusen", that develop beneath the RPE.

Two principal forms, or stages, of AMD have been described. They are referred to as the dry, or atrophic form and the wet, or exudative form. In the dry form, the RPE undergoes necrosis without coincident neovascularization, resulting in atrophy of the RPE. In the much less prevalent wet form, neovascular fronds derived from the choroid invade and detach the RPE, often causing detachment of the neural retina by filling the subretinal space with fluid.[1] The dry form of AMD has been postulated to be an earlier stage of the wet form.[2,3]

AMD appears to be hereditable as an autosomal dominant trait in a significant portion of afflicted individuals.[4,5,6] The delayed onset of AMD, its highly variable phenotypic

manifestations, and its presumed incomplete penetrance present challenges for AMD genetic linkage analyses. Some other inherited maculopathies show phenotypic similarities with AMD and the chromosomal loci of some of these diseases have been determined by linkage analyses. The genes associated with some of these dystrophies have been mapped and in three cases, blue-cone monochromasy,[4] pattern dystrophy,[5,6] and Sorsby's fundus dystrophy,[7] actually identified. For example, Sorsby's fundus dystrophy (SFD) is clinically similar to AMD, but tends to occur at a much earlier age.[8,9] Recent analyses of two families with SFD has identified single point mutations in the TIMP-3 gene located on chromosome 22, where the SFD gene has been mapped.[10] TIMP-3 is an inhibitor of extracellular matrix metalloproteinases, molecules that participate in extracellular matrix turnover. In addition, a mutation in the *bestrophin* gene on chromosome 11 has recently been identified as causing the ocular disorder Best's macular dystrophy.[11] However, none of these genes has been found to be responsible for a significant fraction of typical late-onset macular degeneration.

The biological bases for the photoreceptor and RPE degeneration that characterize MD are poorly understood. Degeneration of both rods and cones may occur concomitant with, or following, RPE degeneration. Photoreceptor cell loss has also been shown to occur in the human macula as a result of normal aging, although the extent and type of cell loss varies significantly.[12,13,14,15,16,17] It has been proposed that a threshold event(s) must occur during normal aging for photoreceptor cell loss to proceed into AMD.[18,19] Factors that have been suggested include light exposure,[19,20,21] decreased choroidal blood flow,[22] decreased density of choroidal blood vessels,[23] RPE degeneration or dysfunction,[15,24,25] and/or the accumulation of abnormal debris and extracellular matrix (ECM) constituents in Bruch's membrane and the choroidal stroma.[26,27] The prevailing dogma is that RPE dysfunction is involved in the earliest stages of AMD, since most of the first clinical and histopathologic signs appear to reside within the RPE and/or Bruch's membrane.[28]

Various explanations for RPE dysfunction have been proposed, including lipofuscin accumulation in the RPE,[18,29,30,31,32,33] changes in the lipid content of Bruch's membrane,[34] the accumulation of extracellular debris in the form of basal laminar deposit and drusen,[35,36,37] macrophage-mediated breakdown of Bruch's membrane,[38] genetic defects,[39,40] or some combination of the above. Genetic defects in extracellular matrix components are expressed in

the eye in a number of diseases,[41,42] for example, an in-frame deletion of the human collagen IX gene (COL9A1) or a deletion of exon 6 in the collagen V gene (COL5a2) can cause eye abnormalities.[41]

The interphotoreceptor matrix (IPM) is an extracellular matrix comprised of an array of proteins, glycoproteins, and proteoglycans, occupying the space between the apical surfaces of the neural retina and the RPE.[43,44,45,46] The IPM is important to the maintenance of normal functions of the neural retina. Based on its strategic location, there is general consensus that the IPM mediates biochemical and physical interactions between the neural retina and RPE, including the exchange of metabolites and catabolic byproducts, regulation of the ionic milieu, mediation of retinal adhesion, sequestration of growth factors, outer segment shedding and phagocytosis, maintenance of photoreceptor polarization, orientation and viability. Retina-RPE adhesion is maintained by a variety of mechanisms including active fluid transport, osmotic pressure gradients, intraocular pressure, physical interaction between RPE and photoreceptor outer segments, microenvironmental conditions and/or components of the IPM. Based on their distribution and composition, IPM glycoconjugates are likely candidates for mediation of the chemical attachment between the neural retina and RPE.[47,48,49,50,51,52,53,54]

Cone matrix sheaths (CMSs), distinct domains of the IPM that contain chondroitin 6-sulfate proteoglycan and surround cone photoreceptor cells, have firm attachments to both RPE cells and the neural retina. This adhesive system is sufficiently strong to detach the RPE or tear the CMS following manual separation of the neural retina from the RPE.[47] The role of IPM constituents in mediating retinal adhesion has been investigated following subretinal or intravitreal injections of various enzymes into rabbit eyes and examination of the consequent morphological, biochemical and physiological changes in retinal structure and function. Chondroitinase ABC, neuraminidase and testicular hyaluronidase, three enzymes that degrade oligosaccharides known to be constituents of the IPM, caused diffuse loss of adhesion that is associated with changes in peanut agglutinin (PNA)-binding to CMSs, without affecting photoreceptor function (based on ERG recordings).[52,53] Enzymatic cleavage of IPM-associated chondroitin sulfate glycosaminoglycans leads to a rapid decrease of retinal adhesiveness in both rabbits and primates *in vivo*.[52,54] In addition, disruption of proteoglycan synthesis *in vivo* leads to loss of CMS-associated chondroitin 6-sulfate proteoglycans, IPM disruption, localized

retinal detachments and photoreceptor outer segment degeneration.[50] Restoration of retinal adhesion, which recovers steadily between 5 and 20 days following exposure to chondroitinase and neuraminidase, correlates closely with the re-establishment of the normal distribution of PNA-binding glycoconjugates in the IPM. Restoration of impaired adhesion is concomitant with the *de novo* biosynthesis of IPM chondroitin sulfate proteoglycans.[53] Thus, it is likely that specific components of the IPM act as major adhesive elements bridging the RPE-retina interface, and that these elements are critically dependent on the metabolic function of RPE and neural retina cells in the microenvironment of the IPM.

Studies in which monkey and human eyes are removed immediately following euthanasia or optic crossclamp, respectively, and the retinas partially "peeled" from the RPE reveal that CMSs remain firmly attached to both the apical RPE and neural retina and elongate up to 4-6 times their normal length in eyes peeled within 45 seconds of enucleation. Additional peeling separation results in separation of the entire RPE from Bruch's membrane, or splitting of the CMSs. These results suggest that adhesion between CMS constituents and the RPE or photoreceptors is stronger than that of the integrity of the CMSs themselves and, as such, provide evidence that CMSs have characteristics consistent with their participation in the establishment and maintenance of retinal attachment.[47,51,52]

Based on the limited amount of information about the roles of proteoglycans in development, it is likely that IPM proteoglycans are involved in the differentiation and maintenance of specific retinal cells and/or with the establishment of physical interactions between photoreceptor and RPE cells. Recent studies have demonstrated that chondroitin sulfate-containing proteoglycans are first detectable in the mouse IPM a few days prior to the elaboration of photoreceptor outer segments. In human retinas, CMS constituents bound by PNA and the chondroitin 6-sulfate antibody (AC6S) are expressed at a time during development when rudimentary cone outer segments first differentiate. Concentrated accumulations of PNA-binding constituents are observed at 17 to 18 weeks. Chondroitin 6-sulfate is first detected in the IPM between 20 and 23 weeks, when rudimentary cone outer segments begin to differentiate and is solely associated with cone outer segments. At this time, cone photoreceptors are well-polarized and enlarged domes of IPM are associated with them. These studies suggest that there is a staggered expression of PNA- and AC6S-

containing IPM constituents, and that the CMSs and IPM constituents may be necessary for subsequent cone outer segment differentiation and survival. This contention is further supported by observations that photoreceptors exhibit some degree of polarity but are unable to maintain differentiated outer segments in culture. In addition, RPE-conditioned medium causes a significant increase in the number of embryonic chick photoreceptor cells forming outer segment-like structures *in vitro* and stimulate their survival and differentiation.

Relatively little is known, however, about the molecular nature, origin, or precise functions of most IPM molecules, especially the heterogeneously distributed aqueous insoluble constituents. Only a limited number of aqueous insoluble IPM constituents have been identified and/or partially characterized. These include: (1) a wheat germ agglutinin (WGA)-binding glycoprotein, designated GP147; (2) a high molecular weight mucin; and (3) a variety of uncharacterized glycoproteins. Early biochemical studies indicated that the "aqueous-extractable" components of the IPM are predominantly proteins and glycoproteins, with some glycosaminoglycans (GAGs), including chondroitin sulfates and hyaluronic acid. The major soluble glycoprotein of the IPM is IRBP, a retinoid-binding protein which has been purified and characterized from a number of species. Other soluble IPM constituents include a variety of enzymes, mucins, complex carbohydrates, S-laminin, and unidentified proteins. Analyses of molecules synthesized and secreted by cultures of RPE and neural retina cells have demonstrated that these cell types are capable of synthesizing and secreting an array of GAGs, including hyaluronate, and proteoglycans. Few of these, however, have been identified as IPM constituents. Furthermore, information pertaining to the cellular source(s) of specific IPM constituents is rare.

The distinct compositional differences between foveal and extrafoveal IPM in adult human retinas, suggests that the foveal IPM may be an outward expression of foveal cone specialization. These cones, which are not fully differentiated until 3-5 years after birth, are structurally distinct from extrafoveal cones in that they more closely resemble rod photoreceptors. Based on our understanding that the viability and functions of cells are dependent upon their interactions with the extracellular matrix, the foveal IPM may be important to the function and homeostasis of cones. It is likely that disruption of the synthesis and/or maintenance of these IPM constituents could result in photoreceptor disease. The

identification of core protein amino acid sequences have allowed investigators to predict potential functions of proteoglycans, which have been tested subsequently *in vivo* and *in vitro*.

Histochemical studies have shown the organization of insoluble IPM glycoconjugates to be complex and to comprise numerous heterologous patterns. These patterns include glycoconjugates exhibiting photoreceptor cell-specific distributions, apical-basal distributions, and uniform distributions. The most well-characterized IPM glycoconjugates are those that display photoreceptor cell-specific distributions. PNA and AC6S antibody react specifically with CMSs. In contrast, rod photoreceptor cell-associated IPM is bound by WGA and *Limax flavus* agglutinin (LFA) which bind to CMSs only weakly. In contrast, molecular constituents of the insoluble IPM have been identified that are selectively localized near the apical surface of the RPE. This domain is specifically labeled by *Phaseolus* agglutinin (PHA-L) and MO-225, an antibody directed against a chondroitin sulfate proteoglycan. This complex picture of IPM heterogeneity is complicated further by observations that show marked differences in the distribution of specific IPM protein glycoconjugates between lower and higher mammals, major light-evoked shifts in the distribution and/or molecular conformation of some constituents, as well as topographical differences (e.g., foveal versus extrafoveal IPM) in the distribution of specific glycoconjugates within the same species.

A number of studies suggest a correlation between changes in IPM composition and the etiology of photoreceptor demise in some degenerations. In the"Royal College of Surgeons" (RCS) rat, or mucopolysaccharidosis VII (MPS VII) mouse, for example, alterations in the distribution of IPM chondroitin 6-sulfate proteoglycan occur prior to photoreceptor degeneration.[55,56,57,58] Similarly, intravitreal injection of β-D-xylopyranoside, a sugar that inhibits GAG chain addition to proteoglycan core proteins, results in cone outer segment degeneration following loss of CMS-associated chondroitin 6-sulfate. These studies strongly indicate that disruption of the normal synthesis or turnover of chondroitin sulfate proteoglycans can lead to cone outer segment degeneration.

The density of rods with abnormal IPM has been shown to increase in correlation with increasing quantities of macular drusen, extracellular deposits in Bruch's membrane associated with AMD. Preliminary analyses of neutral and amino sugars from the IPM of retinas with and without significant numbers of these "PNA-binding rods" show a decrease in IPM sialic

acid concentration. In addition, the densities of PNA-binding rods correlate with drusen grade. The incidence of cone degeneration in these same eyes can be as high as 30-40%, in contrast to age-matched controls without significant macular drusen, where cone photoreceptor loss is approximately 12-15%. These studies provide direct evidence that changes in IPM proteoglycan composition are associated with drusen, and that these changes may be related to photoreceptor dysfunction and visual loss that occurs in individuals with AMD.

The vitronectin receptor (VnR) is associated with apical RPE and cone photoreceptor outer and inner segment plasma membranes. VnR co-localizes with components of apposed CMSs at these cell surfaces, and with actin cables within apical RPE microvilli and cone photoreceptor inner segments. Based on the known functions of VnR, it is likely that its interaction with IPM ligands may modulate other cellular activities, such as translation of external cues into signals that affect cytoskeletal organization or modification of other IPM ligands.

No pharmacologic treatment has been shown to be effective in preventing, arresting, or reversing the loss of vision associated with MD. Based on studies showing strong negative correlations between AMD and dietary plasma carotenoid[59,60,61] or antioxidant levels,[62,63] orally administered antioxidants have been proposed as a measure to reduce AMD incidence and/or severity. Laser photocoagulation has been used with limited success to arrest neovascularization in the "wet" form of the disease. More recently, it has been advanced as a way to facilitate regression of macular drusen in AMD patients, resulting in improved visual acuity, minimal scar formation, and restoration of foveal morphology.[64] Transplantation of retinal pigment epithelial cells or photoreceptor cells as a treatment for retinal degeneration has also been described.[65] Identification of novel IPM components and their genes will make possible novel therapeutic and diagnostic agents for diseases or conditions associated with abnormal IPM, such as retinal detachment, chorioretinal degenerations, retinal degenerations and macular degenerations such as AMD, or other dystrophies or degenerations involving IPM, cones or rods.

SUMMARY OF THE INVENTION

The present invention is based on the discovery of a novel family of mammalian genes and proteins isolated from the interphotoreceptor matrix (IPM). This newly identified gene family is referred to herein as the interphotoreceptor matrix component or "IPMC" gene family and encodes a family of proteoglycans useful as therapeutic and diagnostic agents.

Members of the IPMC gene family have been identified in humans, monkey, cow, goat, rabbit, dog, cat, pig and rat IPM. Two subfamilies of IPMCs are designated the IPM150 family and the IPM200 family, respectively, based upon sequence similarity and homology of the various mammalian cDNAs and proteins to human IPM150 or human IPM200 cDNAs and proteins, respectively. The core proteins of primate IPM150 and IPM200 proteins are unique, based on partial amino acid sequences obtained from the N-terminal and internal peptides, and on deduced amino acid sequences obtained from a PCR product derived from monkey and human retinal cDNA encoding portions of IPM150 and IPM200. Partial amino acid sequences of the N-termini of IPM150 and IPM200 are similar, but exhibit specific inter- and intra-species differences.

The nucleotide sequence of human IPM150 cDNA is 3,261 bp and contains an uninterrupted open reading frame of 2,313 bp, 127 bp of 5' untranslated sequence, and 818 bp of 3' untranslated sequence (Figure 3, SEQ ID NO:3). The IPM150 gene is located on chromosome 6q14.2-q15, which region also contains loci for progressive bifocal chorioretinal atrophy, autosomal dominant Stargardt's-like macular dystrophy, North Carolina macular dystrophy and Salla disease.[66,67,68] The genomic organization and exon/intron boundaries of the IPM150 gene have been characterized.[67] Verification that the cDNA sequence encodes human IPM150 is provided by the presence of the complete amino-terminal amino acid sequence, as well as all internal peptides, within the deduced amino acid sequence. The predicted amino acid sequence revealed a novel protein of 771 amino acids with an isoelectric point of 4.70 and predicted molecular weight of 86.36kDa. The protein is generally hydrophilic except for the first 18 amino acids, which form a hydrophobic region flanked by charged amino acid residues. This region may represent a signal sequence. The IPM150 amino acid sequence features O- and N-linked glycosylation consensus sequences, an EGF-like domain, a single glycosaminoglycan attachment consensus sequence, two clusters of cysteine residues, and two potential hyaluronan binding motifs. Northern blot analysis

suggests that IPM150 mRNA is expressed in the neural retina but not in the RPE or other ocular tissue. *In situ* hybridization analyses indicate that IPM150 mRNA is expressed specifically in cone and rod photoreceptor cells.

IPM200 is a large proteoglycan of the IPMC gene family. Several protein fragments and cDNA clones were isolated which encompass the entire open reading frame of human IPM200. The cDNA sequence of human IPM200 is set forth in SEQ ID NO:5 and Figure 4. The nucleotide sequence of 4,165 bp contains an uninterrupted open reading frame of 3,726 bp, 192 bp of 5' untranslated sequence, and 249 bp of 3' untranslated sequence (Figure 4, SEQ ID NO:5). The predicted amino acid sequence reveals a novel protein of 1,241 amino acids. The IPM200 cDNA sequence features O- and N- linked glycosylation consensus sequences. The IPM200 gene is located on chromosome 3, 5.76cR from the marker WI-3277. IPM200 appears to be the human homolog of PG10.2. an uncharacterized sequence tag expressed in the pineal gland and retina of rats.[69]

In one aspect, the invention features isolated or recombinant IPMC nucleic acid molecules. The IPMC nucleic acids of the invention are useful for identifying additional members of the IPMC family. In one embodiment, the IPMC molecule is expressed in the interphotoreceptor matrix of an individual. In another embodiment, the IPMC molecule is expressed in other organs. The IPMC molecule may be a component of the extracellular matrix of many cell types and may participate in the structure and function of a variety of cells or tissues, e.g., in the maintenance of normal basement membrane. In a preferred embodiment, the nucleic acids hybridize to a nucleic acid having the sequence set forth in SEQ ID NO:1, 3 or 5. In a more preferred embodiment, the nucleic acids have the nucleic acid sequence set forth in SEQ ID NO:1, 3 or 5 or a portion thereof. In one embodiment, the IPMC nucleic acids are from a vertebrate. In a preferred embodiment, the IPMC nucleic acids are from a mammal, e.g., a human. The IPMC nucleic acids can be non-coding, (e.g., a fragment, antisense or ribozyme molecule) or can encode functional IPMC polypeptides (e.g., a polypeptide which functions as either an agonist or antagonist of at least one bioactivity of a native human IPMC protein). In a further embodiment, the IPMC nucleic acids are mutant variants or allelic variants of the nucleic acids shown in SEQ ID NO:1, 3 or 5. In one embodiment, the nucleic acid of the present invention can hybridize to a vertebrate IPMC gene

or to the complement of a vertebrate IPMC gene. In a further embodiment, the claimed nucleic acid can hybridize with a nucleic acid sequence shown in Figures 2, 3 or 4 (SEQ ID NO:1, 3 or 5) or the complement thereof. In a preferred embodiment, the hybridization is conducted under mildly stringent or stringent conditions.

In a preferred embodiment, the IPMC nucleic acid sequence comprises one or more of the following nucleic acid sequences: nucleotides 251 to 769 of SEQ ID NO:3, nucleotides 788 to 1822 of SEQ ID NO:3, nucleotides 1901 to 2018 of SEQ ID NO:3, nucleotides 2189 to 2320 of SEQ ID NO:3 and nucleotides 2330 to 2356 of SEQ ID NO:3. In a preferred embodiment, the IPMC nucleic acid hybridizes with one or more of the following nucleotide sequences: nucleotides 251 to 769 of SEQ ID NO:3, nucleotides 788 to 1822 of SEQ ID NO:3, nucleotides 1901 to 2018 of SEQ ID NO:3, nucleotides 2189 to 2320 of SEQ ID NO:3 and nucleotides 2330 to 2356 of SEQ ID NO:3.

In further embodiments, the nucleic acid molecule is an IPMC nucleic acid that is at least about 70%, preferably about 80%, more preferably about 85%, and even more preferably at least about 90% or 95% homologous to the nucleic acid shown as SEQ ID NO:3 or 5, or to the complement of the nucleic acid shown as SEQ ID NO:3 or 5. In a preferred embodiment, the IPMC nucleic acid comprises at least 10 bp, 50 bp, 100 bp, 200 bp, 300 bp, 400 bp, 500 bp, 600 bp, 700 bp, 800 bp, 900 bp, 1000 bp, 1250 bp, 1500 bp, 1750 bp, 2000 bp, 2500 bp, 3000 bp, 3500 bp, 4000 bp or 4165 bp which share at least 60% identity with or hybridizes to the IPMC nucleic acids of the invention.

The invention also provides probes and primers comprising substantially purified oligonucleotides, which correspond to a region of nucleotide sequence which hybridizes to at least about 6, at least about 10, at least about 15, at least about 20, or preferably at least about 25 consecutive nucleotides of the sequence set forth as SEQ ID NO:3 or 5, or complements of the sequence set forth as SEQ ID NO:3 or 5 or naturally occurring mutants or allelic variants thereof. A list of preferred primers are shown in Table 1, with corresponding SEQ ID NOS: 12-45. In preferred embodiments, the probe or primer further includes a label group attached thereto, which is capable of being detected.

In one aspect of the IPMC of the invention, the subject nucleic acids are linked to an appropriate replicative cloning vector, a number of which are well known in the art. For

expression, the subject nucleic acids can be operably linked to a transcriptional regulatory sequence, e.g., at least one of a transcriptional promoter (e.g., for constitutive expression or inducible expression) or transcriptional enhancer sequence. Such regulatory sequences in conjunction with an IPMC protein nucleic acid molecule can provide a useful vector for gene expression comprising an expression system. This invention also describes host cells transfected with said expression system vector and *in vitro* (e.g., cell culture) and *in vivo* (e.g., transgenic) methods for producing an IPMC protein by employing the expression vectors of the invention.

In another aspect, the invention provides methods for producing IPMC polypeptides and nucleic acids comprising culturing cells under conditions effective for the production of an IPMC polypeptide and harvesting the polypeptide.

In another aspect, the invention features isolated IPMC polypeptides, preferably substantially pure preparations, e.g., of plasma purified or recombinantly produced polypeptides. The IPMC polypeptides can comprise full length proteins or can comprise smaller fragments corresponding to one or more particular motifs/domains, or fragments comprising at least about 5, 10, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 771, 800, 900, 1000, 1100, 1200, or 1241 amino acids in length.

In a preferred embodiment, the IPMC polypeptide is comprised of the amino acid sequence set forth in SEQ ID NO:2, 4 or 6. The subject IPMC polypeptides also include within its scope modified proteins, e.g., proteins which are resistant to post-translational modification, for example, due to mutations which alter modification sites (such as tyrosine, threonine, serine or asparagine residues), or which prevent glycosylation of the protein, or which prevent interaction of the protein with intracellular proteins involved in signal transduction.

The IPMC polypeptides of the invention comprise one or more of domain A (residues 40 to 215 of SEQ ID NO:4), domain B (residues 221 to 565 of SEQ ID NO:4), domain C (residues 591 to 630 of SEQ ID NO:4), and domain D (residues 688 to 731 of SEQ ID NO:4 and/or residues 735 to 743 of SEQ ID NO:4). Any and all nucleotides that encode the above amino acid sequences of domains A, B, C or D are contemplated.

The IPMC polypeptides of the present invention can be glycosylated, or conversely, by choice of the expression system or by modification of the protein sequence to preclude glycosylation, reduced carbohydrate analogs can also be provided. Glycosylated forms can be obtained, for example, based on derivitization with glycosaminoglycan chains.

In yet another preferred embodiment, the invention features a purified or recombinant IPMC polypeptide, which has the ability to modulate, e.g., mimic or antagonize, an activity of a wild-type IPMC protein.

Another aspect of the invention features chimeric molecules (e.g., fusion proteins) comprising an IPMC protein. For instance, the IPMC protein can be provided as a recombinant fusion protein which includes a second polypeptide portion, e.g., a second polypeptide having an amino acid sequence unrelated (heterologous) to the IPMC polypeptide. A preferred IPMC fusion protein is an immunoglobulin-IPMC fusion protein, in which an immunoglobulin constant region is fused to an IPMC polypeptide.

Yet another aspect of the present invention concerns an immunogen comprising an IPMC polypeptide in an immunogenic preparation, the immunogen being capable of eliciting an immune response specific for an IPMC polypeptide; e.g. a humoral response, an antibody response and/or a cellular response. In a preferred embodiment, the immunogen comprises an antigenic determinant, e.g. a unique determinant of a protein encoded by the nucleic acid set forth in SEQ ID NO:3 or 5.

A still further aspect of the present invention features antibodies and other binding proteins or peptides that are specifically reactive with an epitope of an IPMC protein.

The invention also features transgenic non-human animals which include (and preferably express) a heterologous form of an IPMC gene described herein, or which misexpress an endogenous IPMC gene (e.g., an animal in which expression of one or more of the subject IPMC proteins are disrupted). Such transgenic animals can serve as animal models for studying cellular and/or tissue disorders comprising mutated or mis-expressed IPMC alleles or for use in drug screening or testing. Alternatively, such transgenic animals can be useful for expressing recombinant IPMC polypeptides.

The invention further features assays and kits for determining whether an individual's IPMC genes and/or IPMC proteins are defective or deficient (e.g., in activity and/or level),

and/or for determining the identity of IPMC alleles. In one embodiment, the method comprises the step of determining the level of IPMC protein, mRNA and/or the transcription rate of an IPMC gene. In another preferred embodiment, the method comprises detecting, in a tissue of the subject, the presence or absence of a genetic alteration, which is characterized by at least one of the following: a deletion of one or more nucleotides from a gene; an addition of one or more nucleotides to the gene; a substitution of one or more nucleotides of the gene; a gross chromosomal rearrangement of the gene; an alteration in the level of a messenger RNA transcript of the gene; the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; and/or a non-wild type level of the IPMC protein. The assays and kits are useful for detecting susceptibility to a disease or condition associated with abnormal IPMC activity. The assays and kits are also useful in determining the efficacy of a therapeutic regime.

IPMC protein mutations that are particularly likely to cause or contribute to the development of a disease or condition associated with abnormal IPMC activity include mutations which result in the absence of IPMC protein or the synthesis of substantially impaired IPMC proteins. Alternatively, an abnormality in the regulation of an IPMC polypeptide is associated with a disease or disorder associated with abnormal IPMC activity. The abnormal regulation of the IPMC proteins of the invention may be due to alterations in the nucleic acid sequence of the IPMC promoter, 5'UTR or 3'UTR, or may be due to nucleotide alterations within an intron of the IPMC gene. Alterations in the nucleotide sequence flanking the IPMC gene may also result in altered IPMC gene expression. The disease or disorder affected by altered IPMC gene regulation or IPMC gene expression may be an ocular disorder, such as retinal detachment, chorioretinal degeneration, retinal degeneration or macular degeneration, such as age-related macular degeneration. Other diseases potentially associated with abnormal IPM are photoreceptor degenerations, RPE degenerations, mucopolysaccharidoses VII, retinoschisis, trauma, rod-cone dystrophies, cone-rod dystrophies and cone degenerations.

IPMC mutations can be detected by: (i) providing a probe/primer comprised of an oligonucleotide which hybridizes to a sense or antisense sequence of an IPMC gene or naturally occurring mutants thereof, or 5' or 3' flanking sequences naturally associated with

the IPMC gene; (ii) contacting the probe/primer with an appropriate nucleic acid containing sample; and (iii) detecting, by hybridization of the probe/primer to the nucleic acid, the presence or absence of the genetic alteration. The genetic alteration can comprise one or more nucleotide differences. Particularly preferred embodiments comprise: (1) sequencing at least a portion of an IPMC gene, (2) performing a single strand conformation polymorphism (SSCP) analysis to detect differences in electrophoretic mobility between mutant and wild-type nucleic acids; (3) identifying mutant IPMC alleles using spectral genotyping using fluorescent probes or the like; and (4) detecting or quantitating the level of an IPMC protein in an immunoassay using an antibody which is specifically immunoreactive with a wild-type or mutated IPMC protein. In another embodiment, the IPMC proteins of the invention are useful for detecting naturally occurring serum antibodies to IPMC proteins, e.g., in the case of an autoimmune disorder involving auto antibodies to IPMC proteins. Further, the invention is useful for identifying, characterizing and isolating the antigen or epitope to which IPMC auto antibodies bind.

Information obtained using the diagnostic assays described herein (alone or in conjunction with information on another genetic defect, which contributes to the same disease) is useful for diagnosing or confirming that a symptomatic subject has a genetic defect (e.g., in an IPMC gene or in a gene that regulates the expression of an IPMC gene), which causes or contributes to the particular disease or disorder. Alternatively, the information (alone or in conjunction with information on another genetic defect, which contributes to the same disease) can be used prognostically for predicting whether a non-symptomatic subject is likely to develop a disease or condition, which is caused by or contributed to by an abnormal IPMC protein activity or protein level in a subject. In particular, the assays permit one to ascertain an individual's predisposition or increased risk of developing a condition associated with abnormal IPMC activity or a mutation in IPMC proteins, where the mutation is a single nucleotide polymorphism (SNP). Such a diagnostic assay may comprise at least one oligonucleotide primer capable of differentiating between a normal IPMC gene and an IPMC gene with one or more nucleotide differences. Based on the prognostic information, a doctor can recommend a regimen or therapeutic protocol useful for preventing or prolonging onset of a disease or condition associated with abnormal IPMC activity (e.g., an ocular disorder such as

retinal detachment, chorioretinal degeneration, retinal degeneration or macular degeneration) in the individual.

Alternatively, the IPMC proteins may serve as markers for ocular disease or disorders, where these altered expression is a consequence of abnormal regulatory processes such as abnormal production of a heterologous protein capable of altering the nucleic acid or protein synthesis of IPMC proteins, transport of IPMC proteins, binding characteristics of IPMC proteins or other mechanism which may not be a consequence of altered IPMC nucleic acid or protein sequence.

In addition, knowledge of the particular alteration or alterations, resulting in defective or deficient IPMC genes, proteins or their expression in an individual, alone or in conjunction with information on other genetic defects contributing to the same disease (the genetic profile of the particular disease) allows customization of therapy to the individual's genetic profile, the goal of pharmacogenomics. The IPMC genetic profile of individuals in a population is determined and a relationship between the IPMC genetic profiles and the phenotypes of the individuals is determined. For example, an individual's IPMC genetic profile or the genetic profile of a disease or condition associated with abnormal IPMC activity (e.g. macular degeneration) can enable a doctor to: 1) more effectively prescribe a drug that addresses the molecular basis of the disease and 2) better determine the appropriate dosage of a particular drug. For example, the expression level of IPMC proteins, alone or in conjunction with the expression level of other genes known to be involved in macular degeneration, can be measured in many patients at various stages of the disease to generate a transcriptional or protein expression profile of the disease or condition. Expression patterns of individual patients can then be compared to the expression profile of the disease or condition to determine the appropriate drug and dose to administer to the patient.

The ability to target populations expected to show the highest clinical benefit, based on the IPMC protein or disease genetic profile, can enable: 1) the repositioning of marketed drugs with disappointing market results; 2) the rescue of drug candidates whose clinical development has been discontinued as a result of safety or efficacy limitations, which are patient subgroup-specific; and 3) an accelerated and less costly development for drug candidates and more

optimal drug labeling (e.g., since the use of IPMC protein as a marker is useful for optimizing effective dose).

In another aspect, the invention provides methods for identifying a compound which modulates an IPMC protein activity, e.g. the interaction between an IPMC polypeptide and a target peptide or protein. In a preferred embodiment, the method includes the steps of (a) forming a reaction mixture, which includes: (i) an IPMC polypeptide, (ii) an IPMC binding partner and (iii) a test compound; and (b) detecting the interaction of the IPMC polypeptide and the IPMC binding partner. A statistically significant change (potentiation or inhibition) in the interaction of the IPMC polypeptide and IPMC binding partner in the presence of the test compound, relative to the interaction in the absence of the test compound, indicates a potential agonist (mimetic or potentiator) or antagonist (inhibitor) of IPMC bioactivity for the test compound. The reaction mixture can be a cell-free protein preparation, e.g., a reconstituted protein mixture or a cell lysate, or it can be a recombinant cell including a heterologous nucleic acid recombinantly expressing the IPMC binding partner.

In preferred embodiments, the step of detecting interaction of the IPMC protein and IPMC binding partner is a competitive binding assay. In other preferred embodiments, at least one of the IPMC polypeptide and the IPMC binding partner comprises a detectable label, and interaction of the IPMC protein and IPMC binding partner is quantified by detecting the label in the complex. The detectable label can be, e.g., a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. In other embodiments, the complex is detected by an immunoassay. In a further embodiment, the invention provides methods for the isolation and characterization of a molecule capable of binding to an IPMC nucleic acid protein or peptide. In a preferred embodiment the IPMC binding molecule is a nucleic acid, a protein, a glycoconjugate (e.g., glycoprotein or glycolipid), a peptide or a peptide nucleic acid.

Yet another exemplary embodiment provides an assay for screening test compounds to identify agents which modulate the amount of IPMC protein produced by a cell. In one embodiment, the screening assay comprises contacting a cell transfected with a reporter gene operably linked to an IPMC promoter with a test compound and determining the level of expression of the reporter gene. The reporter gene can encode, e.g., a gene product that gives rise to a detectable signal or marker such as: color, fluorescence, luminescence, cell viability,

relief of a cell nutritional requirement, cell growth, and drug resistance. For example, the reporter gene can encode a gene product selected from the group consisting of chloramphenicol acetyl transferase, luciferase, β-galactosidase and alkaline phosphatase.

Also within the scope of the invention are methods for treating or preventing the development of a disease or disorder which is associated with defective or deficient IPMC, comprising administering (e.g., either locally or systemically) to a subject, a pharmaceutically effective amount of a composition comprising an IPMC therapeutic. The method is particularly useful for treating or preventing retinal detachment, chorioretinal degenerations, retinal degenerations or macular degenerations.

Other features and advantages of the invention are apparent from the following detailed description and claims.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a schematic representation of retina and choroid. (A) As seen in histological section, (B) Retinal neurons shown diagrammatically. A, amacrine cells; B, bipolar cells; BM, Bruch's membrane; C, cone cells; CC, choroidocapillaris; ELM, external limiting membrane; G, ganglion cells; GCL, ganglion cell layer; H, horizontal cells; ILM, inner limiting membrane; INL, internal nuclear layer; IPM, interphotoreceptor matrix; IS, inner segments of rods and cones; IPL, internal plexiform layer; NFL, nerve fiber layer; ONL, outer nuclear layer; OPL, outer plexiform layer; OS, outer segments of rods and cones; PE, pigment epithelium; PRL, photoreceptor layer; PT, photorecptor cell terminals; RC, rod cells; ST, stroma vascularis of choroid.

Figure 2 is a DNA sequence of the macaque IPM150 cDNA including the 5' and 3' untranslated regions (UTRs) (SEQ ID NO:1). The predicted amino acid sequence of the open reading frame as provided herein is set forth in SEQ ID NO:2.

Figure 3 is a DNA sequence of the human IPM150 cDNA including the 5' and 3' untranslated regions (UTRs) (SEQ ID NO:3). The predicted amino acid sequence of the open reading frame as provided herein is set forth in SEQ ID NO:4.

Figure 4 is a DNA sequence of the human IPM200 cDNA including the 5' and 3' untranslated regions (UTRs) (SEQ ID NO:5). The predicted amino acid sequence of the open reading frame as provided herein is set forth in SEQ ID NO:6.

~~Figure 5 is an alignment of macaque, pig, human and rat IPMC protein sequences.~~

DETAILED DESCRIPTION OF THE INVENTION

I. General

The present invention is based, at least in part, on the discovery of two novel human genes, termed "hIPM150" and "hIPM200" and a novel macaque gene termed "mIPM150".

Digestion of isolated IPM preparations with chondroitinase ABC released two major glycoproteins which migrate in SDS-polyacrylamide gels with apparent molecular weights of 150 kDa and 200 kDa. These proteoglycans are herein referred to as "IPM150" and "IPM200," respectively and represent two subfamilies of the IPMC family. They appear to comprise a significant portion of the IPM and are bound by both PNA and AC6S antibodies on Western blots of denatured IPM preparations.

Alignment of IPM150 and IPM200 N-terminal amino acid sequences from humans, monkeys and pigs reveals a high degree of sequence conservation between the three species (Figure 5).

RT-PCR of monkey retina total RNA, employing primers designed according to sequence obtained from tryptic digest fragments of macaque IPMC protein repeatedly yielded a 580 bp amplification, designated 70-1 (Figure 2). Clone 70-1 was used as a probe to screen a human retinal cDNA library. Three cDNA clones, designated 8.1.2, 11.1.1 and 12.3.1, were isolated, sequenced and assembled into one contiguous sequence. The assembled nucleotide sequence of 3,261 bp contains an uninterrupted open reading frame of 2,313 bp, 127 bp of untranslated 5' sequence and 818 bp of untranslated 3' sequence (Figure 3). Verification that the assembled sequence encodes human IPM150 is provided by the presence of the complete amino-terminal amino acid sequence, as well as all internal peptides, within the deduced amino acid sequence.

Clone 8.1.2 lacks a 234 bp segment within its 5' region that encodes the amino terminus of IPM150. In an attempt to determine if this clone represented a cloning artifact or

an IPM150 isoform and to demonstrate that the isolated clones in fact represent the IPM150 cDNA, a PCR product encompassing the entire open reading frame of IPM150 was cloned. PCR amplification of reverse-transcribed human retinal RNA yields a 2,579 bp fragment which was subcloned and designated 9p#3. This clone is completely homologous to the assembled sequence and includes the 234 bp region that is absent in clone 8.1.2.

NCBI database searches of the compiled nucleotide and amino acid sequences show that they are novel. A few expressed sequence tags (ESTs), derived from human retina, vein endothelial cells, and brain share some homology with IPM150 (GenBank Accession numbers H38604, W26960, H38594, AA326863 and AA296278). Translation of the open reading frame of the assembled IPM150 cDNA sequences encodes a protein of 771 amino acids with an isoelectric point of 4.70 and a predicted molecular weight of 86.36 kDa. The protein is generally hydrophilic[70] except for the first 18 amino acids which form a hydrophobic region flanked by charged amino acid residues. This region may represent a signal sequence.

A distinct distribution of consensus sequence sites for N- and O-linked glycosylation are present in human and monkey IPM150 and IPM200. There are four possible consensus sequences for N-linked glycosylation[71] sequestered in the amino-terminus (amino acid residues 42, 143, 191 and 215) and four probable consensus sequences in the carboxy-terminus (amino acid residues 592, 616, 630 and 648) portions of the human IPM150 core protein. In contrast, the central domain of IPM150 (between amino acid residues 221 and 565) features 16 sites which are suitable for O-linked glycosylation, as predicted by a proposed algorithm for the activity of polypeptide N-acetylgalactosaminyl-transferase (amino acid residues 221, 222, 407, 433, 435, 441, 442, 443, 445, 470, 497, 527, 531, 537, 561 and 565).[72] There is almost no overlap between the regions containing potential O- and N-linked glycosylation sites. Hyaluronan binding consensus sequences[73] (amino acid residues 735-743) and several cysteine residues are also present in the amino- and carboxy-terminal regions of the core protein. Further, a glycosaminogycan (GAG) attachment site is located at amino acid residue 296. The distribution of the carboxy-terminal cysteine residues closely resembles that of EGF-like domains[74] (amino acid residues 688-731), motifs that are present in many extracellular matrix proteins. One reported function of the domain is to promote the survival of neighboring cells. By analogy, it is conceivable that this region of IPM150 may promote photoreceptor viability

in vivo. This contention is supported by previous studies suggesting that the IPM, and more specifically its proteoglycans, are important in maintaining photoreceptor cell viability.[50,75]

The IPM200 cDNA sequence also encodes nineteen O-linked glycosylation sites (amino acid residues 183, 191, 400, 408, 430, 439, 482, 496, 533, 537, 556, 624, 718, 730, 801, 813, 815, 822, and 870) and six N-linked glycosylation sites (amino acid residues 154, 301, 320, 370, 942 and 956). The IPM200 cDNA encodes a GAG attachment site at amino acid residue 604.

The predicted molecular weight of the "native" IPM150 core protein, as deduced from its cDNA sequence, is about 77.8 kDa. The protein migrates at approximately 150 kDa on SDS-PAGE gels under denaturing conditions following digestion of native IPM with chondroitinase ABC. The discrepancy between the calculated and the observed molecular weights of IPM150 is most likely due to the presence of N and/or O-glycosidically linked oligosaccharides that remain attached to the protein after chondroitinase treatment. This contention is supported by biochemical analyses demonstrating that IPM150 migrates at 80 kDa and 105 kDa following enzymatic or chemical deglycosylation, respectively.

With the exception of the amino-terminal signal sequence, which appears to be removed during the maturation of IPM150, the core protein does not possess any other hydrophobic regions that would constitute transmembrane domains. This implies that IPM150 is secreted into the interphotoreceptor space and is not membrane intercalated. Based on previous studies supporting a role for IPM proteoglycans in adhesion of the neural retina to the RPE,[47,48,53,54] IPM150 may act as a bridging element, between as yet unidentified membrane-associated molecules on the photoreceptor and RPE cell surfaces to effect retinal adhesion. For example, IPM150 may bind to the vitronectin receptor, which has recently been localized to tile surfaces of photoreceptor outer segments and the apical microvilli of RPE cells. The chondroitin-sulfate glycosaminoglycans of IPM150 might also participate in retinal adhesion, through receptors like CD44,[76] an adhesion molecule associated with Müller cell apical microvilli in mice and humans.[43]

Based on studies suggesting that hyaluronan is an IPM component[43,77,78,79] and that IPMC proteins possess two hyaluronan binding domains, IPM150 may also interact with hyaluronan to effect retinal adhesion. The N-terminus of IPM150 contains two potential

hyaluronan-binding motifs. Previous biochemical studies have indicated that hyaluronic acid is present as a component of the IPM and that hyaluronidase disrupts CMSs *in vitro* and weakens retinal adhesion *in vivo*. Hyaluronan may stabilize the IPM through interactions with CD44, IPM150, IPM200 and perhaps other insoluble IPM constituents. It is likely that these are as yet unidentified IPM protein ligands. Such a system would provide a framework for a molecular model of retinal adhesion, whereby IPM molecules, such as insoluble CMS-associated proteoglycans, bind to the extracellular domains of VnR and CD44. The molecules of the present invention are useful for elucidating the distribution and spatial relationships between IPM150, IPM200, CD44, VnR and hyaluronan.

IPM150 isoforms or spliced variants, which are synthesized by rod photoreceptors and do not contain chondroitin 6-sulfate, are also anticipated. For example, these isoforms could be produced by differential glycosylation and/or small differences in the amino acid sequences of rod and cone IPM150. The absence of 234 nucleotides in clone 8.1.2 suggests that splicing isoforms may indeed exist, although the difference in overall length of the transcripts is too small to lead to distinguishable bands on Northern blot analyses.

Biochemical studies have clearly documented that chondroitin 6-sulfate moities are associated with IPM150.[44] Earlier biochemical studies have documented that chondroitin 6-sulfate moities are associated with IPM150.[44] The deduced amino acid sequence for IPM150 does not contain the typical glycosaminoglycan attachment consensus sequence SGXG.[80] However, a single attachment site identical to those observed in collagen type IX or chick decorin is present in the central portion of the protein.[81,82] Alternatively, spliced variants or isoforms of IPM150 that possess Ser-Gly dipeptides may also exist.

In order to determine the cellular source(s) of IPM150, Northern and *in situ* hybridization analyses were conducted. IPM150 cDNA probes hybridize to a 3.9 kb transcript that is present in relatively high abundance in retinal RNA. Occasionally, a larger transcript of approximately 6.5 kb is also detected, albeit at a much reduced signal strength. No signal is detected on Northern blots of RNA isolated from RPE/choroid, iris or cornea. Distinct hybridization of IPM150 antisense riboprobes to the human retinal outer nuclear layer (ONL) was observed on sections of human retina, RPE and choroid. IPM150 transcripts are present

within both rod and cone photoreceptor cells. No labeling of any other region of the eye was observed.

IPM150, or a related molecule, is transcribed in non-ocular tissues. Dot-blot analyses of mRNA derived from a variety of human tissues indicate that IPM150-derived probes hybridize to human adult and fetal lung and thymus, as well as to adult kidney and small intestine. In addition, ESTs partially homologous to the IPM150 cDNA sequence have been identified from human endothelial cell and brain cDNA libraries. The extraocular biosynthesis of IPMC mRNA suggests that IPMC proteins function generally as extracellular matrix proteins in a number of tissues.

The 4,165 bp nucleotide sequence of the human IPM200 cDNA contains an open reading frame of 3,726 bp, 192 bp of 5' untranslated sequence, and 249 bp of 3' untranslated sequence (Figure 4, SEQ ID NO:5).

The IPMC nucleic acid sequences of the present inventions enable the development of specific probes and the generation of transgenic animal nucleic acid models to more closely examine the function of this unique proteoglycan *in vivo*.

II. Definitions

For convenience, the meaning of certain terms and phrases employed in the specification, examples, and appended claims are provided below.

The term "an aberrant activity", as applied to an activity of a polypeptide such as an IPMC protein, refers to an activity which differs from the activity of the wild-type or native polypeptide or which differs from the activity of the polypeptide in a healthy subject. An activity of a polypeptide can be aberrant because it is stronger than the activity of its native counterpart. Alternatively, an activity can be aberrant because it is weaker or absent relative to the activity of its native counterpart. An aberrant activity can also be a change in an activity or in its binding characteristics. For example an aberrant polypeptide can interact with a different target peptide. A cell can have an aberrant IPMC protein activity due to overexpression or underexpression of the gene encoding an IPMC protein.

The terms "abnormal" or "abnormality", as used herein, are meant to refer to a condition which differs from that found normally. Such abnormality when referring to a

sequence means a change in the sequence such as an addition, deletion or change (e.g., substitution of nucleic acid or amino acid for another nucleic acid or amino acid, respectively). For example, such abnormality when referring to protein activity means a change in the expression level of RNA or protein, RNA or protein turnover, activity or binding characteristics. Such abnormality when referring to a medical condition means a difference in the structure, function or composition of an organ, system, cell or organism.

The term "agonist", as used herein, is meant to refer to an agent that mimics or upregulates (e.g. potentiates or supplements) an IPMC bioactivity. An IPMC agonist can be a wild-type IPMC protein or derivative thereof having at least one bioactivity of a wild-type IPMC protein. An "IPMC therapeutic" can also be a compound that upregulates or down regulates expression of an IPMC gene or which increases at least one bioactivity of an IPMC protein. An agonist can also be a compound which increases the interaction of an IPMC polypeptide with another molecule, e.g, an upstream region of a gene, which is regulated by an IPMC transcription factor.

The term "allele", which is used interchangeably herein with "allelic variant" refers to alternative forms of a gene or portions thereof. Alleles occupy the same locus or position on homologous chromosomes. When a subject has two identical alleles of a gene, the subject is said to be homozygous for the gene or allele. When a subject has two different alleles of a gene, the subject is said to be heterozygous for the gene. Alleles of a specific gene can differ from each other in a single nucleotide, or several nucleotides, and can include substitutions, deletions, and insertions of nucleotides. An allele of a gene can also be a form of a gene containing a mutation. The term "allelic variant of a polymorphic region of an IPMC gene" refers to a region of an IPMC gene having one or several nucleotide sequences found in that region of the gene in other individuals.

The term "antagonist" as used herein is meant to refer to an agent that downregulates (e.g. suppresses or inhibits) at least one IPMC bioactivity. An IPMC antagonist can be a compound which inhibits or decreases the interaction between an IPMC protein and another molecule, e.g, an upstream region of a gene, which is regulated by an IPMC transcription factor. Accordingly, a preferred antagonist is a compound which inhibits or decreases binding to an upstream region of a gene, which is regulated by an IPMC transcription factor and

thereby blocks subsequent activation of the IPMC protein. An antagonist can also be a compound that downregulates expression of an IPMC gene or which reduces the amount of IPMC protein present, e.g., by decreasing protein synthesis or increasing protein turnover. The IPMC antagonist can be a dominant negative form of an IPMC polypeptide, e.g., a form of an IPMC polypeptide which is capable of interacting with an upstream region of a gene, which is regulated by an IPMC transcription factor, but which is not capable of regulating transcription. The IPMC antagonist can also be a nucleic acid encoding a dominant negative form of an IPMC polypeptide, an IPMC antisense nucleic acid, or a ribozyme capable of interacting specifically with an IPMC RNA. Yet other IPMC antagonists are molecules which bind to an IPMC polypeptide or its receptor and inhibit its action. Such molecules include peptides, antibodies and small molecules.

"Biological activity" or "bioactivity" or "activity" or "biological function", which are used interchangeably, for the purposes herein means an effector or antigenic function that is directly or indirectly performed by an IPMC polypeptide (whether in its native or denatured conformation), or by any subsequence thereof. Biological activities include binding to a target nucleic acid e.g, an upstream region of a gene, which is regulated by an IPMC transcription factor. An IPMC bioactivity can be modulated by directly affecting an IPMC polypeptide. Alternatively, an IPMC bioactivity can be modulated by modulating the level of an IPMC polypeptide, such as by modulating expression of an IPMC gene or by modulating the turnover of the IPMC protein.

As used herein the term "bioactive fragment of an IPMC polypeptide" refers to a fragment of a full-length IPMC polypeptide, wherein the fragment specifically mimics or antagonizes the activity of a wild-type IPMC polypeptide. The bioactive fragment preferably is a fragment capable of interacting with e.g., an upstream region of a gene, which is regulated by an IPMC transcription factor.

"Cells", "host cells" or "recombinant host cells" are terms used interchangeably herein, and refer to cells which are capable of or have been transformed with a vector, typically an expression vector. A host cell can be prokaryotic or eukaryotic, including bacteria, insect, yeast and mammalian cells. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain

modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A "chimeric polypeptide" or "fusion polypeptide" is a fusion of a first amino acid sequence encoding one of the subject IPMC polypeptides with a second amino acid sequence defining a domain (e.g. polypeptide portion) foreign to and not substantially homologous with any domain of an IPMC polypeptide. A chimeric polypeptide may present a foreign domain which is found (albeit in a different polypeptide) in an organism which also expresses the first polypeptide, or it may be an "interspecies", "intergenic", etc. fusion of polypeptide structures expressed by different kinds of organisms. In general, a fusion polypeptide can be represented by the general formula X-IPMC-Y, wherein the IPMC protein represents a portion of the polypeptide which is derived from an IPMC polypeptide, and X and Y are independently absent or represent amino acid sequences which are not related to an IPMC protein sequence in an organism, including naturally occurring mutants.

The term “nucleotide sequence complementary to the nucleotide sequence set forth in SEQ ID NO:x” refers to the nucleotide sequence of the complementary strand of a nucleic acid strand having SEQ ID NO:x. The term “complementary strand” is used herein interchangeably with the term “complement”. The complement of a nucleic acid strand can be the complement of a coding strand or the complement of a non-coding strand. When referring to double stranded nucleic acids, the complement of a nucleic acid having SEQ ID NO:x refers to the complementary strand of the strand having SEQ ID NO:x or to any nucleic acid having the nucleotide sequence of the complementary strand of SEQ ID NO:x. When referring to a single stranded nucleic acid having the nucleotide sequence SEQ ID NO:x, the complement of this nucleic acid is a nucleic acid having a nucleotide sequence which is complementary to that of SEQ ID NO:x. The nucleotide sequences and complementary sequences thereof are always given in the 5' to 3' direction.

A "delivery complex" shall mean a targeting means (e.g. a molecule that results in higher affinity binding of a gene, protein, polypeptide or peptide to a target cell surface and/or increased cellular or nuclear uptake by a target cell). Examples of targeting means include: sterols (e.g. cholesterol), lipids (e.g. a cationic lipid, virosome or liposome), viruses (e.g.

adenovirus, adeno-associated virus, and retrovirus) or target cell specific binding agents (e.g. ligands recognized by target cell specific receptors). Preferred complexes are sufficiently stable *in vivo* to prevent significant uncoupling prior to internalization by the target cell. However, the complex is cleavable under appropriate conditions within the cell so that the gene, protein, polypeptide or peptide is released in a functional form.

As is well known, genes may exist in single or multiple copies within the genome of an individual. Such duplicate genes may be identical or may have certain modifications, including nucleotide substitutions, additions or deletions, which all still code for polypeptides having substantially the same activity. The term "DNA sequence encoding an IPMC polypeptide" may refer to one or more genes within a particular individual. Moreover, certain differences in nucleotide sequences may exist between individual organisms, which are called alleles. Such allelic differences may or may not result in differences in amino acid sequence of the encoded polypeptide, yet still encode a polypeptide with the same biological activity.

The term "disease or condition associated with abnormal IPMC activity" is a condition in which an IPMC molecule is abnormally expressed. The IPMC molecule may abnormally be overexpressed or underexpressed, for example. The activity of the IPMC molecule may be abnormal in that its functional activity may be increased or decreased. The activity of the IPMC molecule may be abnormal by having changed its function. Such a change in function may be a consequence of a change in the nucleic acid sequence or a change in its protein sequence.

The term "extracellular matrix" ("ECM") refers to the collagens, proteoglycans, non-collagenous glycoproteins and elastins that may surround cells and provide structural support for cells as well as maintain various functions of cells, such as cell adhesion, proliferation and differentiation. See, for example,[41,42,83]. A skilled artisan will appreciate that the precise composition and physical properties of ECM, as well as its function, varies between various cell types and between various organs.

The term "hybridizes" refers to the annealing of one nucleic acid sequence to another. Appropriate stringency conditions which promote DNA hybridization, for example, 2 to 6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C, are known to those skilled in the art or can be found in *Current Protocols in Molecular*

Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. The salt concentration in the wash step can be selected from a low stringency of about 6.0 x SSC to a high stringency of about 0.1 x SSC. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Formamide may be added to the hybridization steps and washing steps in order to decrease the temperature requirement by 1°C per 1% formamide added.

The terms "IPMC polypeptide" and "IPMC protein" are interchangeable as used herein, and intended to encompass the family comprising polypeptides or proteins having the amino acid sequences shown as SEQ ID NO:2, 4 or 6, or fragments thereof, mutant and allelic variants, and homologs thereof. In an embodiment, the IPMC polypeptides have the amino acid sequence of SEQ ID NO:2, 4 or 6.

The term "IPM150 nucleic acid" refers to a nucleic acid encoding an IPM150 protein, such as nucleic acids having SEQ ID NO:1 or 3, as well as fragments thereof, complements thereof, mutant and allelic variants, homologs, and derivatives thereof. The terms "IPM150 polypeptide" and "IPM150 protein" are interchangeable as used herein, and are intended to encompass polypeptides comprising the amino acid sequence shown as SEQ ID NOS:2 or 4, or fragments thereof, mutant and allelic variants, and homologs thereof and include agonist and antagonist polypeptides.

The term "IPM200 nucleic acid" refers to a nucleic acid encoding an IPM200 protein, such as nucleic acids having SEQ ID NO:5, as well as fragments thereof, complements thereof, mutant and allelic variants, homologs, and derivatives thereof. The terms "IPM200 polypeptide" and "IPM200 protein" are interchangeable and intended to encompass polypeptides comprising the amino acid sequence shown as SEQ ID NO:6 or fragments thereof, mutant and allelic variants, and homologs thereof and include agonist and antagonist polypeptides.

"Homology" or "identity" or "similarity" refers to sequence similarity between two peptides or between two nucleic acid molecules and can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base or amino acid, then the molecules are identical at that position. A degree of homology or similarity or identity between nucleic acid

sequences is a function of the number of identical or matching nucleotides at positions shared by the nucleic acid sequences. An "unrelated" or "non-homologous" sequence shares less than about 40% identity, though preferably less than about 25 % identity, with one of the IPM protein sequences of the present invention. % identity, homology or similarity are determined using a sequence alignment software as described herein. Alternatively % identity, homology or similarity are determined by the number of nucleotide or amino acid differences in a sequence of a certain length. For example, a 100 residue sequence with 20 residue differences is defined as 80% identical, wherein a difference means a different residue or lack of residue. "Homologous" refers to the evolutionary relatedness of two nucleic acid or protein sequences. "Identity" refers to the degree to which nucleic acids or amino acids are the same between two sequences. "Similarity" refers to the degree to which nucleic acids or amino acids are the same, but includes neutral amino acid substitutions that do not significantly change the function of the protein as is well known in the art. Similarity also refers to neutral degenerate nucleic acids that may be substituted within a codon without changing the amino acid identity of the codon, as is well known in the art.

The term "interact" as used herein is meant to include detectable relationships or associations (e.g. biochemical interactions) between molecules, such as interaction between protein-protein, protein-nucleic acid, nucleic acid-nucleic acid, and protein-small molecule or nucleic acid-small molecule in nature.

The term "isolated" as used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules separated from other DNAs, or RNAs, respectively, that are present in the natural source of the macromolecule. For example, an isolated nucleic acid encoding one of the subject IPMC polypeptides preferably includes no more than about 10 kilobases (kb) of nucleic acid sequence which naturally immediately flanks the IPMC gene in genomic DNA, more preferably no more than about 5kb of such naturally occurring flanking sequences, and most preferably about 1.5kb of such naturally occurring flanking sequence. The term isolated as used herein also refers to a nucleic acid or peptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally

occurring as fragments and would not be found in the natural state. The term "isolated" is also used herein to refer to polypeptides which are isolated from other cellular proteins and is meant to encompass both purified and recombinant polypeptides.

The term "macular degeneration" refers to a condition in which the macula degenerates, e.g., as a consequence of decreased growth of cells of the macula, increased death or rearrangement of the cells of the macula, or a combination of these events. Macular degeneration results in the loss of integrity of the histoarchitecture of the cells of the normal macula and/or the loss of function of the cells of the macula. Any condition which alters or damages the integrity or function of the macula may be considered to fall within the definition of macular degeneration. Other examples of diseases in which cellular degeneration has been implicated include retinal detachment, chorioretinal degenerations, retinal degenerations, photoreceptor degenerations, RPE degenerations, mucopolysaccharidoses VII, rod-cone dystrophies, cone-rod dystrophies and cone degenerations.

The term "modulation" as used herein refers to both upregulation (i.e., activation or stimulation (e.g., by agonizing or potentiating)) and downregulation (i.e. inhibition or suppression (e.g., by antagonizing, decreasing or inhibiting)).

The term "mutated gene" refers to an allelic form of a gene, which is capable of altering the phenotype of a subject having the mutated gene relative to a subject which does not have the mutated gene. If a subject must be homozygous for this mutation to have an altered phenotype, the mutation is said to be recessive. If one copy of the mutated gene is sufficient to alter the genotype of the subject, the mutation is said to be dominant. If a subject has one copy of the mutated gene and has a phenotype that is intermediate between that of a homozygous and that of a heterozygous subject (for that gene), the mutation is said to be co-dominant.

The "non-human animals" of the invention include mammals such as rodents, non-human primates, sheep, dog, cow, pig, chickens, amphibians, reptiles, etc. Preferred non-human animals are selected from the primate family (e.g., macaque) or rodent family including rat and mouse, most preferably rat, though transgenic amphibians, such as members of the *Xenopus* genus, and transgenic chickens can also provide important tools for understanding and identifying agents which can affect, for example, embryogenesis and tissue formation and

cellular differentiation. The term "chimeric animal" is used herein to refer to animals in which the recombinant gene is found, or in which the recombinant gene is expressed in some but not all cells of the animal. The term "tissue-specific chimeric animal" indicates that one of the recombinant IPMC genes is present and/or expressed or disrupted in some tissues but not others.

As used herein, the term "nucleic acid" refers to polynucleotides or oligonucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs and as applicable to the embodiment being described, single (sense or antisense) and double-stranded polynucleotides.

The term “polymorphism” refers to the coexistence of more than one form of a gene or portion (e.g., allelic variant) thereof. A portion of a gene of which there are at least two different forms, i.e., two different nucleotide sequences, is referred to as a “polymorphic region of a gene”. A polymorphic region can be a single nucleotide, the identity of which differs in different alleles. A polymorphic region can also be several nucleotides long. A “polymorphic gene” refers to a gene having at least one polymorphic region.

As used herein, the term "promoter" refers to a DNA sequence that regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in cells. The term encompasses "tissue specific" promoters, i.e. promoters, which effect expression of the selected DNA sequence only in specific cells (e.g. cells of a specific tissue). The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well. The term also encompasses non-tissue specific promoters and promoters that constitutively express or that are inducible (i.e. expression levels can be controlled).

The terms "protein", "polypeptide" and "peptide" are used interchangeably herein when referring to a gene product. The term "recombinant protein" refers to a polypeptide of the present invention which is produced by recombinant DNA techniques, wherein generally, DNA encoding an IPMC polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. Likewise the term “recombinant nucleic acid” or “recombinant DNA” refers to a nucleic acid or DNA of the

present invention which is produced by recombinant DNA techniques, wherein generally, DNA encoding an IPMC polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. Moreover, the phrase "derived from", with respect to a recombinant IPMC gene, is meant to include within the meaning of "recombinant protein" those proteins having an amino acid sequence of a native IPM polypeptide, or an amino acid sequence similar thereto which is generated by mutations including substitutions and deletions (including truncation) of a naturally occurring form of the polypeptide.

"Small molecule" as used herein, is meant to refer to a composition, which has a molecular weight of less than about 5 kD and most preferably less than about 4 kD. Small molecules can be nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic (carbon containing) or inorganic molecules. Many pharmaceutical companies have extensive libraries of chemical and/or biological mixtures, often fungal, bacterial, or algal extracts, which can be screened with any of the assays of the invention to identify compounds that modulate an IPMC bioactivity.

As used herein, the term "specifically hybridizes" or "specifically detects" refers to the ability of a nucleic acid molecule of the invention to hybridize to at least approximately 6, 12, 20, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 1000, 1500, 2000, 2500, 3000, 4000, or 4165 consecutive nucleotides of a vertebrate gene, preferably an IPMC gene.

The term "IPMC therapeutic" refers to various forms of IPMC polypeptides, as well as peptidomimetics, nucleic acids, or small molecules, which can modulate at least one activity of an IPMC polypeptide, e.g., binding to and/or otherwise regulating expression of a gene, by mimicking or potentiating (agonizing) or inhibiting (antagonizing) the effects of a naturally-occurring IPMC polypeptide. An IPMC therapeutic which mimics or potentiates the activity of a wild-type IPMC polypeptide is an "IPMC agonist". Conversely, an IPMC therapeutic which inhibits the activity of a wild-type IPMC polypeptide is an "IPMC antagonist".

The term "transcriptional regulatory sequence" is a generic term used throughout the specification to refer to DNA sequences, such as initiation signals, enhancers, and promoters, which induce or control transcription of protein coding sequences with which they are operably linked. In preferred embodiments, transcription of one of the IPMC genes is under

the control of a promoter sequence (or other transcriptional regulatory sequence) which controls the expression of the recombinant gene in a cell-type in which expression is intended. It is understood that the recombinant gene can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally-occurring forms of a IPMC polypeptide.

As used herein, the term "transfection" means the introduction of a nucleic acid, e.g., via an expression vector, into a recipient cell by nucleic acid-mediated gene transfer. "Transformation", as used herein, refers to a process in which a cell's genotype is changed as a result of the cellular uptake of exogenous DNA or RNA, and, for example, the transformed cell expresses a recombinant form of an IPMC polypeptide or, in the case of anti-sense expression from the transferred gene, the expression of a naturally-occurring form of the IPMC polypeptide is disrupted.

As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., one of the IPMC polypeptides, or an antisense transcript thereto) which has been introduced into a cell. A transgene could be partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, can be homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can also be present in a cell in the form of an episome. A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as 5' UTR sequences, 3' UTR sequences, or introns, that may be necessary for optimal expression of a selected nucleic acid.

A "transgenic animal" refers to any animal, preferably a non-human mammal, bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a

recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. In the typical transgenic animals described herein, the transgene causes cells to express a recombinant form of one of the IPMC polypeptides, e.g. either agonistic or antagonistic forms. However, transgenic animals in which the recombinant IPMC gene is silent are also contemplated, as for example, the FLP or CRE recombinase dependent constructs described below. Moreover, "transgenic animal" also includes those recombinant animals in which gene disruption of one or more IPMC genes is caused by human intervention, including both recombination and antisense techniques.

The term "treating" as used herein is intended to encompass curing as well as ameliorating at least one symptom of the condition or disease.

The terms "vector," "cloning vector," or "replicative cloning vector," are interchangeable as used herein, and refer to a nucleic acid molecule, which is capable of transporting another nucleic acid to which it has been linked. One type of preferred vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors." The term "expression system" as used herein refers to an expression vector under conditions whereby an mRNA may be transcribed and/or an mRNA may be translated into protein. The expression system may be an *in vitro* expression system, which is commercially available or readily made according to art known techniques, or may be an *in vivo* expression system, such as a eukaryotic or prokaryotic cell containing the expression vector. In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. In the present specification, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

The term "wild-type allele" refers to an allele of a gene which, when present in two copies in a subject results in a wild-type phenotype. There can be several different wild-type

alleles of a specific gene, since certain nucleotide changes in a gene may not affect the phenotype of a subject having two copies of the gene with the nucleotide changes.

III. Nucleic Acids of the Present Invention

The invention provides IPMC nucleic acids, mutant ar allelic variants, homologs thereof, and portions thereof. Preferred nucleic acids comprise a sequence, which is at least about 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, and more preferably 85% homologous with a nucleotide sequence of an IPMC gene, e.g., such as a sequence shown in one of SEQ ID NOS:1, 3 or 5, or complements thereof. Nucleic acids at least 90%, more preferably 95%, and most preferably at least about 98-99% homologous with a nucleic sequence represented in one of SEQ ID NOS:1, 3 or 5, or a complement thereof are of course also within the scope of the invention.

The invention also pertains to isolated nucleic acids comprising a nucleotide sequence encoding IPM polypeptides, mutant or allelic variants and/or equivalents of such nucleic acids. The term equivalent is understood to include nucleotide sequences encoding functionally equivalent IPM polypeptides or functionally equivalent peptides having an activity of an IPMC protein such as described herein. Equivalent nucleotide sequences include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and therefore includes sequences that differ from the nucleotide sequence of the IPMC cDNA shown in one of SEQ ID NOS:1, 3 or 5, due to the degeneracy of the genetic code.

Preferred nucleic acids are vertebrate IPMC nucleic acids. Particularly preferred vertebrate IPMC nucleic acids are mammalian. Regardless of species, particularly preferred IPMC nucleic acids encode polypeptides that are at least about 60%, 65%, 70%, 72%, 74%, 76%, 78%, 80%, 90%, or 95% similar or identical to an amino acid sequence of a vertebrate IPMC protein. In one embodiment, the nucleic acid is a cDNA encoding a polypeptide having at least one bio-activity of the subject IPMC polypeptide. Preferably, the nucleic acid includes all or a portion of the nucleotide sequence corresponding to the nucleic acid of SEQ ID NO:1, 3 or 5. Useful algorithms include the Needleman-Wunsch-Sellers algorithm, the Smith-Waterman algorithm, the Lipman-Pearson algorithm, and the Kaslin-Altschul algorithm, for

example. A number of algorithm-based software packages are available for analyzing and comparing nucleic acid and protein sequences, including, for example, GCG, BESTFIT and BLAST. Exemplary parameters for using the BESTFIT program have a gap weight of about 1 to about 99 and a length weight of about .01 to about 10. A gap weight of 50 and length weight of 3 are considered the default parameters for the BLAST program. For example, a BLAST alignment of portions of the human IPM150 and monkey IPM150 sequences (amino acid residues 73-256 and 2-185, respectively), using a gap weight of 12 and a length weight of 4 revealed that the amino acid sequences share 95% similarity and 90.8% identity within the aligned regions. A BLAST alignment of human IPM150 and monkey IPM150 cDNA sequences (nucleotides 344-894 and 4-554, respectively), using a gap weight of 50 and a length weight of 3 revealed that the nucleotide sequences share 95.5% sequence identity in the aligned regions. A BLAST alignment of the human IPM150 cDNA and the human IPM200 cDNA (nucleotides 12-3260 and 3-4042, respectively) using a gap weight of 30 and a length weight of 1 revealed that the nucleotide sequences share 73.4% sequence identity in the aligned regions. A BLAST alignment of the human IPM150 cDNA and the human IPM200 cDNA (nucleotides 1839-2274 and 2881-3316, respectively) using a gap weight of 50 and a length weight of 3 revealed that the nucleotide sequences share 68.3% sequence identity in the aligned regions.

Still other preferred nucleic acids of the present invention encode an IPM polypeptide which is comprised of at least 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 or 1241 amino acid residues. For example, such nucleic acids can comprise about 150, 300, 450, 600, 750, 900, 1050, 1200, 1350, 1500, 1800, 2100, 2400, 2700, 3000 or 3723 base pairs. Also within the scope of the invention are nucleic acid molecules for use as probes/primer or antisense molecules (i.e. noncoding nucleic acid molecules), which can comprise at least about 6, 12, 20, 30, 50, 60, 70, 80, 90, 100, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000 or 4165 base pairs in length.

Another aspect of the invention provides a nucleic acid which hybridizes to a nucleic acid represented by SEQ ID NOS:1, 3 or 5, or a complement thereof. Appropriate stringency conditions which promote DNA hybridization, for example, 2.0 (high stringency) to 6.0 (low stringency) x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x

SSC at 50°C, are known to those skilled in the art or can be found in Current Protocols in Molecular Biology (1989), *John Wiley & Sons, N.Y.* 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 6.0 x SSC to a high stringency of about 0.1 x SSC. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Further, the addition of formamide to the hybridization conditions increases the stringency and allows hybridization at lower temperatures, approximately by 1°C cooler for every 1% formamide added. The temperature, salt, or formamide content may be varied, or temperature and salt concentration may be held constant while the other variable is changed. In a preferred embodiment, an IPMC nucleic acid of the present invention binds to one of SEQ ID NOS:1, 3 or 5, or complement thereof under moderately stringent conditions, for example at about 2.0 x SSC and about 40°C. In a particularly preferred embodiment, an IPMC nucleic acid of the present invention binds to one of SEQ ID NO:1, 3 or 5, or a complement thereof under high stringency conditions.

Nucleic acids having a sequence that differs from the nucleotide sequences shown in one of SEQ ID NOS:1, 3 or 5, or a complement thereof due to degeneracy in the genetic code are also within the scope of the invention. Such nucleic acids encode functionally equivalent peptides (i.e., peptides having a biological activity of an IPMC polypeptide) but differ in sequence from the sequence shown in the sequence listing due to degeneracy in the genetic code. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC each encode histidine) may result in "silent" mutations which do not affect the amino acid sequence of an IPMC polypeptide. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the subject IPMC polypeptides exist among mammals. One skilled in the art will appreciate that these variations in one or more nucleotides (e.g., up to about 3-5% of the nucleotides) of the nucleic acids encoding polypeptides having an activity of an IPMC polypeptide may exist among individuals of a given species due to natural allelic variation.

The polynucleotide of the present invention may also be fused in frame to a marker sequence, also referred to herein as "Tag sequence" encoding a "Tag peptide", which allows

for marking and/or purification of the polypeptide of the present invention. In a preferred embodiment, the marker sequence is a hexahistidine tag, e.g., supplied by a PQE-9 vector. Numerous other Tag peptides are available commercially. Other frequently used Tags include myc-epitopes[84] which includes a 10-residue sequence from c-myc, the pFLAG system (International Biotechnologies, Inc.), the pEZZ-protein A system (Pharmacia, NJ), and a 16 amino acid portion of the *Haemophilus influenza* hemagglutinin protein. Furthermore, any polypeptide can be used as a Tag so long as a reagent, e.g., an antibody interacting specifically with the Tag polypeptide is available or can be prepared or identified.

In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence at the N-terminus of the desired portion of the recombinant protein, can allow purification of the expressed fusion protein by affinity chromatography using a Ni^{2+} metal resin. The purification leader sequence can then be subsequently removed by treatment with enterokinase to provide the purified protein.[85,86]

Techniques for making fusion genes are known to those skilled in the art. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed to generate a chimeric gene sequence.[87]

Other preferred IPMC fusion proteins include IPMC-Ig fusion proteins, such as, e.g., IPM150-immunoglobulin or IPM200-immunoglobulin (IPM150-Ig or IPM200-Ig) polypeptides. An IPM150-Ig or IPM200-Ig polypeptide can comprise the entire extracellular domain of IPMC protein, e.g, human IPMC protein, or a variant thereof. For example, an IPM150-Ig or IPM200-Ig fusion proteins can be prepared as described e.g., in U.S. Patent No. 5,434,131.

As indicated by the examples set out below, IPMC nucleic acids can be obtained from mRNA present in any of a number of eukaryotic cells, e.g., from eye tissue. It should also be possible to obtain nucleic acids encoding IPMC polypeptides of the present invention from genomic DNA from any nucleated cell. For example, a gene encoding an IPMC protein can be cloned from either a cDNA or a genomic library in accordance with protocols described herein, as well as those generally known to persons skilled in the art. cDNA encoding an IPMC protein can be obtained by isolating total mRNA from a cell, e.g., a vertebrate cell, a mammalian cell, or a human cell, including embryonic cells. Double stranded cDNAs can then be prepared from the total mRNA, and subsequently inserted into a suitable plasmid or bacteriophage vector using any one of a number of known techniques. The gene encoding an IPMC protein can also be cloned using established polymerase chain reaction techniques in accordance with the nucleotide sequence information provided by the invention. The nucleic acid of the invention can be DNA or RNA or analogs thereof. A preferred nucleic acid is a cDNA represented by a sequence selected from the group consisting of SEQ ID NOS:1, 3 or 5.

Preferred nucleic acids encode a vertebrate IPM polypeptide comprising an amino acid sequence that is at least about 60% homologous, more preferably at least about 70% homologous and most preferably at least about 80% homologous with an amino acid sequence set forth in SEQ ID NO:3 or 5. Nucleic acids which encode polypeptides with at least about 90%, more preferably at least about 95%, and most preferably at least about 98-99% homology with an amino acid sequence set forth in SEQ ID NO:3 or 5 are also within the scope of the invention. In one embodiment, the nucleic acid is a cDNA encoding a peptide having at least one activity of the subject vertebrate IPMC polypeptide. Preferably, the nucleic acid includes all or a portion of the nucleotide sequence corresponding to the coding region of SEQ ID NO:1, 3 or 5.

Preferred nucleic acids encode a bioactive fragment of a vertebrate IPMC polypeptide comprising an amino acid sequence, which is at least about 60% homologous or identical, more preferably at least about 70% homologous or identical, still more preferably at least about 75% homologous or identical and most preferably at least about 80% homologous or identical with an amino acid sequence of SEQ ID NO:3 or 5. Nucleic acids which encode polypeptides which are at least about 90%, more preferably at least about 95%, and most

preferably at least about 98-99% homologous or identical, with an amino acid sequence represented in SEQ ID NO:4 or 6 are also within the scope of the invention.

Bioactive fragments of IPMC polypeptides can be polypeptides, which bind upstream of and/or regulate the expression of a gene. The bioactive fragment may be glycosylated. Assays for determining whether an IPMC polypeptide has any of these or other biological activities are known in the art and are further described herein.

Nucleic acids encoding modified forms or mutant forms of IPMC proteins also include those encoding IPMC proteins having mutated glycosylation sites, such that either the encoded IPMC protein is not glycosylated, partially glycosylated and/or has a modified glycosylation pattern. Other preferred nucleic acids of the invention include nucleic acids encoding derivatives of IPMC polypeptides which lack one or more biological activities of IPMC polypeptides. Such nucleic acids can be obtained, e.g., by a first round of screening of libraries for the presence or absence of a first activity and a second round of screening for the presence or absence of another activity.

Also within the scope of the invention are nucleic acids encoding splice variants or nucleic acids representing transcripts synthesized from an alternative transcriptional initiation site, such as those whose transcription was initiated from a site in an intron.

In preferred embodiments, the IPMC nucleic acids can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids.[88] As used herein, the term "peptide nucleic acids" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone. The neutral backbone of a peptide nucleic acid has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of peptide nucleic acid oligomers can be performed using standard solid phase peptide synthesis protocols as described in reference 84.[89]

Peptide nucleic acids of IPMC proteins can be used in therapeutic and diagnostic applications and are further described herein. Such modified nucleic acids can be used as antisense or antigene agents for sequence-specific modulation of gene expression or in the

analysis of single base pair mutations in a gene by, e.g., peptide nucleic acids directed PCR clamping or as probes or primers for DNA sequence and hybridization.[88,89]

Peptide nucleic acids of IPMC protein can further be modified, e.g., to enhance their stability or cellular uptake, e.g., by attaching lipophilic or other helper groups to the IPMC peptide nucleic acids, by the formation of peptide nucleic acids-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. IPMC peptide nucleic acids can also be linked to DNA as described.[88,90] For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the peptide nucleic acids and the 5' end of DNA.[91] Peptide nucleic acids monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' peptide nucleic acids segment and a 3' DNA segment.[90] Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' peptide nucleic acids segment.[92]

In other embodiments, IPMC nucleic acids may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo),* or agents that facilitate transport across the cell membrane as described herein.

IV. <u>Probes and Primers</u>

The nucleotide sequences described herein further allow for the generation of probes and primers designed for use in identifying and/or cloning IPMC homologs in other cell types, e.g., from other tissues, as well as IPMC homologs from other mammalian organisms. For instance, the present invention also provides a probe/primer comprising a substantially purified oligonucleotide, which oligonucleotide comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least approximately 12, preferably 25, more preferably 40, 50 or 75 consecutive nucleotides of sense or anti-sense sequence selected from the group consisting of SEQ ID NOS:3 or 5, or naturally occurring mutants thereof. For instance, primers based on the nucleic acid sequence represented in SEQ ID NO:3 or 5 can be used in PCR reactions to clone IPMC homologs. Exemplary primer pairs and PCR conditions are described below.

IPM 150 Primer Pairs (SEQ ID NOs:12-45)	Annealing Temp.	cDNA Product (bp)
IPM-18-F GGATTTTTCTCCAAGTTCAAGG IPM-18-R ACGGGGGTTAAAGTCTGTCC	62°C	208
IPM-17-F CGAACAAAAAGATCCGCATT IPM-17-R CCTTCTGCCTCTTTGACATTG	56°C	222
IPM-16-F ATCAGGACTGGGTCAGCATC IPM-16-R TCAACAGATGTTGCCAACGT	58°C	210
IPM-15-F GAGCCTGGTGAAACCATTGT IPM-15-R GTGGAGCTCAGCGTCTCTCT	60°C	194
IPM-14-F TGTGTTGGAGGAGCAGAGG IPM-14-R AAAAGATGGCTCAAGCTCCA	58°C	213
IPM-13-F GAAACTTCCAGGATTCAAAAAA IPM-13-R AGGAGGACAAGCAACCAGAA	58°C	217
IPM-12-F TCCAACAAAATTGAAAGTGAGG IPM-12-R AATCAGAGCTGCCCACATCT	60°C	217
IPM-11-F AGCCTTTGGTCCTGACACC IPM-11-R CCACCTTTCTTTATGGCATCA	60°C	210
IPM-10-F AGTGCAGGTGGCGAAGATAT IPM-10-R CTCCCTGTCAGAAGCTCCAC	60°C	219
IPM-9-F CCACCTGCATCTTCAGATGACA IPM-9-R AGTTCTATGACCATTGCCCC	60°C	209
IPM-8-F AGGATACCACTCCTGTCTCAGC IPM-8-R AACAATTCACACAGCTGCTGG	62°C	186
IPM-7-F CCGAGCTCTGGAGCAAC IPM-7-R GAGGATTTTCGTTCTGCTGC	56°C	199
IPM-6-F TCTGTGCCGTATAACCTCAC IPM-6-R GACTGAGGAAGCGGAGTGTC	60°C	202
IPM-5-F ACGAACGGACTGAGGAAG IPM-5-R TTCTGAATTACTGACCGTAGAA	56°C	245
IPM-4-F TTCCAAAATCAACAAAATAACA IPM-4-R GGTCATCAAAATCCAGACATA	54°C	207
IPM-3-F TGCCTTCTCAAGGAAAATGGAGACAGG IPM-3-R TAAGCCAGGTTTGCTTCCAC	80°C 60°C	242
IPM-2-F TAAAACCCCAAATGCAATCA IPM-2-R GCAGGTCTCTCTAAACGCATG	54°C	244

Likewise, probes based on the subject IPMC protein sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins, for use, e.g., in prognostic or diagnostic assays. In preferred embodiments, the probe further comprises a label group attached thereto and able to be detected, e.g., the label group can be selected, for example, from among radioisotopes, fluorescent compounds, enzymes and enzyme co-factors well known in the art.

V. Antisense, Ribozyme and Triplex Techniques

Another aspect of the invention relates to the use of the isolated nucleic acids in antisense therapy. As used herein, "antisense" therapy refers to administration or *in situ* generation of oligonucleotide molecules or their derivatives which specifically hybridize (e.g., bind) under cellular conditions, with the cellular mRNA and/or genomic DNA encoding one or more of the subject IPMC proteins so as to inhibit expression of that protein, e.g., by inhibiting transcription and/or translation. The binding may be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, "antisense" therapy refers to the range of techniques generally employed in the art, and includes any therapy which relies on specific binding to oligonucleotide sequences.

An antisense construct of the present invention can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the cellular mRNA which encodes an IPMC protein. Alternatively, the antisense construct can be an oligonucleotide probe which is generated *ex vivo* and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences of an IPMC gene. Such oligonucleotide probes are preferably modified oligonucleotides which are resistant to endogenous nucleases, e.g., exonucleases and/or endonucleases, and are therefore stable *in vivo*. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patent Nos. 5,176,996, 5,264,564 and 5,256,775). Additionally, general approaches to

constructing oligomers useful in antisense therapy have been reviewed.[93,94] With respect to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, e.g., between the -10 and +10 regions of the IPMC nucleotide sequence of interest, are preferred.

Antisense approaches involve the design of oligonucleotides (either DNA or RNA) that are complementary to IPMC mRNA. The antisense oligonucleotides bind to the IPMC mRNA transcripts and prevent translation or promote degradation of the transcript. Absolute complementarity, although preferred, is not required. In the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize depends on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the mRNA, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the coding and 3' untranslated sequences of mRNAs have been shown to be effective at inhibiting translation of mRNAs as well.[95] Therefore, oligonucleotides complementary to either the 5' or 3' untranslated, non-coding regions of an IPMC gene could be used in an antisense approach to inhibit translation of endogenous IPMC mRNA. While antisense nucleotides complementary to the IPMC coding region sequence can be used, those complementary to the transcribed untranslated region and to the region comprising the initiating methionine are most preferred. Oligonucleotides complementary to the 5' untranslated region of the mRNA should preferably include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could also be used in accordance with the invention. Whether designed to hybridize to the 5', 3' or coding region of IPMC mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably less than about 100 and more preferably less than about 50, 25, 17 or 10 nucleotides in length.

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors), or agents facilitating transport across the cell membrane[96,97,98] or the blood-brain barrier,[99] hybridization-triggered cleavage agents[100] or intercalating agents.[101] To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxytiethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-

oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

The antisense oligonucleotide can also comprise peptide nucleic acid oligomers, as described above.[89,102] One advantage of peptide nucleic acid-oligomers is their ability to bind to complementary DNA essentially independently from the ionic strength of the medium due to the neutral backbone of the DNA. In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet a further embodiment, the antisense oligonucleotide is an α-anomeric oligonucleotide. An α-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other.[103] The oligonucleotide can be a 2'-0-methylribonucleotide,[104] or a chimeric RNA-DNA analogue.[105]

Oligonucleotides of the invention may be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized,[106] methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports,[107] etc.

The antisense molecules can be delivered to cells which express IPMC protein *in vivo*. A number of methods have been developed for delivering antisense DNA or RNA to cells, e.g., antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systematically.

However, it may be difficult to achieve intracellular concentrations of the antisense sufficient to suppress translation on endogenous mRNAs in certain instances. Therefore a preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells in the patient results in the transcription of sufficient amounts of single stranded RNAs that forms complementary base pairs with the endogenous IPMC transcripts and thereby prevent translation of the IPMC mRNA. For example, a vector can be introduced *in vivo* such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive and can include but not be limited to: the SV40 early promoter region,[108] the promoter contained in the 3' long terminal repeat of Rous sarcoma virus,[109] the herpes thymidine kinase promoter,[110] the regulatory sequences of the metallothionein gene,[111] etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct which can be introduced directly into the tissue site. Alternatively, viral vectors can be used which selectively infect the desired tissue, in which case administration may be accomplished by another route (e.g., systematically).

Ribozyme molecules designed to catalytically cleave IPMC mRNA transcripts can also be used to prevent translation of IPMC mRNA and expression of IPMC protein.[112,113,114] While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy IPMC mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: [5'-UG-3']. The construction and production of hammerhead ribozymes[115] is well known in the art. There are a number of potential hammerhead ribozyme cleavage sites within the nucleotide sequence of human IPMC cDNA. Preferably the

ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the IPMC mRNA; i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention can also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators.[116,117,118,119,120] The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in an IPMC gene.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.) and should be delivered to cells which express the IPMC gene *in vivo*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells produce sufficient quantities of the ribozyme to destroy endogenous IPMC messages and inhibit translation. Because ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Endogenous IPMC gene expression can also be reduced by inactivating or "knocking out" the IPMC gene or its promoter using targeted homologous recombination.[121,122,123] For example, a mutant, non-functional IPMC protein (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous IPMC gene (either the coding regions or regulatory regions of the IPMC gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express IPMC protein *in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the IPMC gene. Such approaches are particularly suited in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive IPMC protein.[122,123] However this approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors.

Alternatively, endogenous IPMC gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the IPMC gene (i.e., the IPMC promoter and/or enhancers) to form triple helical structures that prevent transcription of the IPMC gene in target cells in the body.[124,125,126]

Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription are preferably single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides should promote triple helix formation via Hoogsteen base pairing rules, which generally require sizable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which result in TAT and CGC triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, containing a stretch of G residues. These molecules form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in CGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizable stretch of either purines or pyrimidines to be present on one strand of a duplex.

Antisense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense

cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Moreover, various well-known modifications to nucleic acid molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

VI. Vectors Encoding IPMC Proteins and IPMC Protein Expressing Cells

The invention further provides plasmids and vectors encoding an IPMC protein, which can be used to express an IPMC protein in a host cell. The host cell may be any prokaryotic or eukaryotic cell. Thus, a nucleotide sequence derived from the cloning of mammalian IPMC proteins, encoding all or a selected portion of the full-length protein, can be used to produce a recombinant form of an IPM polypeptide via microbial or eukaryotic cellular processes. Ligating the polynucleotide sequence into a gene construct, such as an expression vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial) cells, are standard procedures well known in the art.

Vectors that allow expression of a nucleic acid in a cell are referred to as expression vectors. Typically, expression vectors used for expressing an IPMC protein contain a nucleic acid encoding an IPMC polypeptide, operably linked to at least one transcriptional regulatory sequence. Regulatory sequences are art-recognized and are selected to direct expression of the subject IPMC proteins. Transcription regulatory sequences have been described.[127] In one embodiment, the expression vector includes a recombinant gene encoding a peptide having an agonistic activity of a subject IPMC polypeptide, or alternatively, encoding a peptide which is an antagonistic form of an IPMC protein.

Suitable vectors for the expression of an IPMC polypeptide include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli*.

A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae.*[128] These vectors can replicate in *E. coli* due the presence of the pBR322 *ori*, and in S. *cerevisiae* due to the replication determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin can be used. In an illustrative embodiment, an IPMC polypeptide is produced recombinantly utilizing an expression vector generated by sub-cloning the coding sequence of one of the IPMC protein genes represented in SEQ ID NOS:3 or 5.

The preferred mammalian expression vectors contain both prokaryotic sequences, to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNAI/amp, pcDNAI/neo, pRc/CMV, pSV2gpt, pSV2neo, pSV2-dhfr, pTk2, pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papillomavirus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures.[129,130]

In some instances, it may be desirable to express the recombinant IPMC polypeptide by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the ß-gal containing pBlueBac III)

When it is desirable to express only a portion of an IPMC protein, such as a form lacking a portion of the N-terminus, i.e. a truncation mutant which lacks the signal peptide, it may be necessary to add a start codon (ATG) to the oligonucleotide fragment containing the desired sequence to be expressed. It is well known in the art that a methionine at the N-

terminal position can be enzymatically cleaved by the use of the enzyme methionine aminopeptidase (MAP). MAP has been cloned from *E. coli*[130] and *Salmonella typhimurium* and its *in vitro* activity has been demonstrated on recombinant proteins.[131] Therefore, removal of an N-terminal methionine, if desired, can be achieved either *in vivo* by expressing IPMC protein derived polypeptides in a host which produces MAP (e.g., *E. coli* or CM89 or *S. cerevisiae*), or *in vitro* by use of purified MAP.[131]

Moreover, the gene constructs of the present invention can also be used as part of a gene therapy protocol to deliver nucleic acids encoding either an agonistic or antagonistic form of one of the subject IPMC proteins. Thus, another aspect of the invention features expression vectors for *in vivo* or *in vitro* transfection and expression of an IPM polypeptide in particular cell types so as to reconstitute the function of, or alternatively, abrogate the function of IPMC protein in a tissue. This could be desirable, for example, when the naturally-occurring form of the protein is misexpressed or the natural protein is mutated and less active.

In addition to viral transfer methods, non-viral methods can also be employed to cause expression of a subject IPM polypeptide in the tissue of an animal. Most nonviral methods of gene transfer rely on normal mechanisms used by mammalian cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral targeting means of the present invention rely on endocytic pathways for the uptake of the subject IPM polypeptide gene by the targeted cell. Exemplary targeting means of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes.

In other embodiments, transgenic animals, described in more detail below could be used to produce recombinant proteins.

VII. Polypeptides of the Present Invention

The present invention makes available IPMC polypeptides which are isolated from, or otherwise substantially free of other cellular proteins. The term "substantially free of other cellular proteins" (also referred to herein as "contaminating proteins") or "substantially pure or purified preparations" are defined as encompassing preparations of IPMC polypeptides having less than about 20% (by dry weight) contaminating protein, and preferably having less than

about 5% contaminating protein. Functional forms of the subject polypeptides can be prepared, for the first time, as purified preparations by using a cloned gene as described herein.

Preferred IPMC proteins of the invention have an amino acid sequence which is at least about 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95% or 96% identical or homologous to an amino acid sequence of SEQ ID NO:4 or 6. Even more preferred IPMC proteins comprise an amino acid sequence which is at least about 97, 98, or 99% homologous or identical to an amino acid sequence of SEQ ID NO:4 or 6. Such proteins can be recombinant proteins, and can be, e.g., produced *in vitro* from nucleic acids comprising a nucleotide sequence set forth in SEQ ID NO:4 or 6, or homologs thereof. For example, recombinant polypeptides preferred by the present invention can be encoded by a nucleic acid, which is at least about 85% homologous and more preferably at least about 90% homologous and most preferably at least about 95% homologous with a nucleotide sequence set forth in SEQ ID NO:4 or 6. Polypeptides which are encoded by a nucleic acid that is at least about 98-99% homologous with the sequence of SEQ ID NO:4 or 6 are also within the scope of the invention.

In a preferred embodiment, an IPMC protein of the present invention is a mammalian IPMC protein. In a particularly preferred embodiment an IPM150 is set forth as SEQ ID NO:4 and IPM200 protein is set forth as SEQ ID NO:6. In particularly preferred embodiments, an IPM150 protein has an IPM150 bioactivity and an IPM200 protein has an IPM200 bioactivity. It is understood that certain post-translational modifications, e.g., phosphorylation, glycosylation and the like, can increase the apparent molecular weight of the IPMC protein relative to the unmodified polypeptide chain.

The invention also features protein isoforms encoded by splice variants of the present invention. Such isoforms may have biological activities identical to or different from those possessed by the IPMC proteins specified by SEQ ID NO:4 or 6.

IPMC polypeptides preferably are capable of functioning as either an agonist or antagonist of at least one biological activity of a wild-type ("authentic") IPMC protein of the appended sequence listing. The term "evolutionarily related to", with respect to amino acid sequences of IPMC proteins, refers to both polypeptides having amino acid sequences which

have arisen naturally, and also to mutational variants of human IPM polypeptides which are derived, for example, by combinatorial mutagenesis.

Full length cDNAs or proteins or fragments corresponding to one or more particular motifs and/or domains or to arbitrary sizes, for example, at least 5, 10, 25, 50, 75, 100, 250, 500, 750, 1000 or 1241, amino acids in length that share greater than 60% amino acid or nucleotide sequence similarity are within the scope of the present invention. See, for example, an alignment of the human IPM150 and monkey IPM150 amino termini in Figure 5.

For example, isolated IPMC polypeptides can be encoded by all or a portion of a nucleic acid sequence shown in any of SEQ ID NOS:1, 3 or 5. Isolated peptidyl portions of IPMC proteins can be obtained by screening peptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such peptides. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, an IPMC polypeptide of the present invention may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of a wild-type (e.g., "authentic") IPMC protein.

Preferred IPMC polypeptides contain one or more of four domains with similarity to four domains of the IPM150 protein, namely domain A, B, C or D. Domain A is located in the amino terminus of the protein, from residue 42 to 215 of SEQ ID NO:4 (Figure 3) containing four possible glycosylation sites (amino acids 42, 143, 191 and 215). Domain B is a central domain containing 17 possible O-linked glycosylation sites from residue 221 to 565 of SEQ ID NO:4 (amino acids 221, 222, 406, 433, 435, 441, 442, 443, 445, 470, 497, 527, 529, 537, 561 and 565). Domain C is located in the carboxy terminus of the protein, from residue 591 to 630 of SEQ ID NO:4 and contains four probable N-linked glycosylation sites (amino acids 591, 592, 616 and 630). Domain D is located in the carboxy-terminus of the protein and contains a hyaluronan binding consensus sequence at residues 735 to 743 of SEQ ID NO:4 and an EGF-like domain between amino acids 688 to 731 of SEQ ID NO:4.

In general, polypeptides referred to herein as having an IPMC activity (e.g., are "bioactive") are defined as polypeptides which include an amino acid sequence encoded by all or a portion of the nucleic acid sequences shown in one of SEQ ID NOS:1, 3 or 5, and which mimic or antagonize all or a portion of the biological/biochemical activities of a naturally occurring IPMC protein. Furthermore, these fragments can either promote or inhibit these processes or agonize or antagonize the activity of another agent which itself promotes or inhibits these processes. Other biological activities of the subject IPMC proteins are reasonably apparent to one of skill in the art. According to the present invention, a polypeptide has biological activity if it is a specific agonist or antagonist of a naturally-occurring form of an IPMC protein. Assays for determining whether a compound, e.g, a protein, such as an IPMC protein or variant thereof, has one or more of the above biological activities are well known in the art.

Other preferred proteins of the invention are those encoded by the nucleic acids set forth in the section pertaining to nucleic acids of the invention. In particular, the invention provides fusion proteins, e.g., IPMC-immunoglobulin fusion proteins. Such fusion proteins can provide, e.g., enhanced stability and solubility of IPMC proteins and may thus be useful in therapy. Fusion proteins can also be used to produce an immunogenic fragment of an IPMC protein. For example, the VP6 capsid protein of rotavirus can be used as an immunologic carrier protein for portions of the IPMC polypeptide, either in the monomeric form or in the form of a viral particle. The nucleic acid sequences corresponding to the portion of a subject IPMC protein to which antibodies are to be raised can be incorporated into a fusion gene construct which includes coding sequences for a late vaccinia virus structural protein to produce a set of recombinant viruses expressing fusion proteins comprising IPMC protein epitopes as part of the virion. It has been demonstrated with the use of immunogenic fusion proteins utilizing the Hepatitis B surface antigen fusion proteins that recombinant Hepatitis B virions can be utilized in this role as well. Similarly, chimeric constructs coding for fusion proteins containing a portion of an IPMC protein and the poliovirus capsid protein can be created to enhance immunogenicity of the set of polypeptide antigens.[132,133,134,135]

The Multiple Antigen Peptide system for peptide-based immunization can also be utilized to generate an immunogen, wherein a desired portion of an IPMC polypeptide is

obtained directly from organo-chemical synthesis of the peptide onto an oligomeric branching lysine core.[136,137] Antigenic determinants of IPMC proteins can also be expressed and presented by bacterial cells.

In addition to utilizing fusion proteins to enhance immunogenicity, it is widely appreciated that fusion proteins can also facilitate the expression of proteins, and accordingly, can be used in the expression of the IPMC polypeptides of the present invention. For example, IPMC polypeptides can be generated as glutathione-S-transferase (GST-fusion) proteins. Such GST-fusion proteins can enable easy purification of the IPMC polypeptide, as for example by the use of glutathione-derivatized matrices.[87]

The present invention further pertains to methods of producing the subject IPMC polypeptides. For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding the subject polypeptides can be cultured under appropriate conditions to allow expression of the peptide to occur. Suitable media for cell culture are well known in the art. The recombinant IPMC polypeptide can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such peptides. In a preferred embodiment, the recombinant IPMC polypeptide is a fusion protein containing a domain which facilitates its purification, such as GST fusion protein.

Moreover, it can be generally appreciated that, under certain circumstances, it may be advantageous to provide homologs of one of the subject IPMC polypeptides, which function in a limited capacity as one of either an IPMC protein agonist (mimetic) or an IPMC antagonist, in order to promote or inhibit only a subset of the biological activities of the naturally-occurring form of the protein. Thus, specific biological effects can be elicited by treatment with a homolog of limited function, and with fewer side effects relative to treatment with agonists or antagonists which are directed to all of the biological activities of naturally occurring forms of IPMC proteins. Homologs of each of the subject IPMC proteins can be generated by mutagenesis, such as by discrete point mutation(s), or by truncation. For instance, mutation can give rise to homologs which retain substantially the same, or merely a subset, of the biological activity of the IPMC polypeptide from which it was derived.

Alternatively, antagonistic forms of the protein can be generated which are able to inhibit the function of the naturally occurring form of the protein, such as by competitively binding to an IPMC receptor.

The recombinant IPMC polypeptides of the present invention also include homologs of the wildtype IPMC proteins, such as versions of those protein which are resistant to proteolytic cleavage, as for example, due to mutations which alter ubiquitination or other enzymatic targeting associated with the protein.

IPMC polypeptides may also be chemically modified to create IPMC protein derivatives by forming covalent or aggregate conjugates with other chemical moieties, such as glycosyl groups, lipids, phosphate, acetyl groups and the like. Covalent derivatives of IPMC proteins can be prepared by linking the chemical moieties to functional groups on amino acid side chains of the protein or at the N-terminus or at the C-terminus of the polypeptide.

Modification of the structure of the subject IPMC polypeptides can be for such purposes as enhancing therapeutic or prophylactic efficacy, stability (e.g., *ex vivo* shelf life and resistance to proteolytic degradation), or post-translational modifications (e.g., to alter phosphorylation pattern of protein). Such modified peptides, when designed to retain at least one activity of the naturally-occurring form of the protein, or to produce specific antagonists thereof, are considered functional equivalents of the IPMC polypeptides described in more detail herein. Such modified peptides can be produced, for instance, by amino acid substitution, deletion, or addition. The substitutional variant may be a substituted conserved amino acid or a substituted non-conserved amino acid.

For example, it is reasonable to expect that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (i.e. isosteric and/or isoelectric mutations) do not have a major effect on the biological activity of the resulting molecule. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) nonpolar = alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar = glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. In

similar fashion, the amino acid repertoire can be grouped as (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine histidine, (3) aliphatic = glycine, alanine, valine, leucine, isoleucine, serine, threonine, with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic = phenylalanine, tyrosine, tryptophan; (5) amide = asparagine, glutamine; and (6) sulfur -containing = cysteine and methionine.[138] Whether a change in the amino acid sequence of a peptide results in a functional IPMC homolog (e.g., functional in the sense that the resulting polypeptide mimics or antagonizes the wild-type form) can be readily determined by assessing the ability of the variant peptide to produce a response in cells in a fashion similar to the wild-type protein, or competitively inhibit such a response. Polypeptides in which more than one replacement has taken place can readily be tested in the same manner.

This invention further contemplates a method for generating sets of combinatorial mutants of the subject IPMC proteins as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g., homologs). The purpose of screening such combinatorial libraries is to generate, for example, novel IPMC homologs which can act as either agonists or antagonist, or alternatively, possess novel activities all together. Thus, combinatorially-derived homologs can be generated to have an increased potency relative to a naturally occurring form of the protein.

In one embodiment, the variegated library of IPMC variants is generated by combinatorial mutagenesis at the nucleic acid level, and is encoded by a variegated gene library. For instance, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential IPMC sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of IPMC sequences therein.

There are many ways by which such libraries of potential IPMC homologs can be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential IPMC sequences. The synthesis of degenerate oligonucleotides is well known in the

art.[139,140,141,142,143,144] Such techniques have been employed in the directed evolution of other proteins.[145,146,147,148,149,151]

Likewise, a library of coding sequence fragments can be provided for an IPMC clone in order to generate a variegated population of IPMC fragments for screening and subsequent selection of bioactive fragments. A variety of techniques are known in the art for generating such libraries, including chemical synthesis. In one embodiment, a library of coding sequence fragments can be generated by (i) treating a double stranded PCR fragment of an IPMC coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule; (ii) denaturing the double stranded DNA; (iii) renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products; (iv) removing single stranded portions from reformed duplexes by treatment with S1 nuclease; and (v) ligating the resulting fragment library into an expression vector. By this exemplary method, an expression library can be derived which codes for N-terminal, C-terminal and internal fragments of various sizes.

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a certain property. Such techniques are generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of IPMC homologs. The most widely used techniques for screening large gene libraries typically comprises cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays described below are amenable to high through-put analysis as necessary to screen large numbers of degenerate IPMC sequences created by combinatorial mutagenesis techniques. Combinatorial mutagenesis has a potential to generate very large libraries of mutant proteins, e.g., in the order of 10^{26} molecules. Combinatorial libraries of this size may be technically challenging to screen even with high throughput screening assays. To overcome this problem, a new technique has been developed recently, recrusive ensemble mutagenesis (REM), which allows one to avoid the very high proportion of non-functional proteins in a random library

and simply enhances the frequency of functional proteins, thus decreasing the complexity required to achieve a useful sampling of sequence space. REM is an algorithm which enhances the frequency of functional mutants in a library when an appropriate selection or screening method is employed.[152,153,154,155]

The invention also provides for reduction of the IPMC proteins to generate mimetics, e.g., peptide or non-peptide agents, such as small molecules, which are able to disrupt binding of an IPMC polypeptide of the present invention with a molecule, e.g. target peptide. Thus, such mutagenic techniques as described above are also useful to map the determinants of the IPMC proteins which participate in protein-protein interactions involved in, for example, binding of the subject IPMC polypeptide to a target peptide. To illustrate, the critical residues of a subject IPMC polypeptide which are involved in molecular recognition of its receptor can be determined and used to generate IPMC protein derived peptidomimetics or small molecules which competitively inhibit binding of the authentic IPMC protein with that moiety. By employing, for example, scanning mutagenesis to map the amino acid residues of the subject IPMC proteins which are involved in binding other proteins, peptidomimetic compounds can be generated which mimic those residues of the IPMC protein which facilitate the interaction. Such mimetics may then be used to interfere with the normal function of an IPMC protein. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine,[156] azepine,[157] substituted gamma lactam rings,[158] keto-methylene pseudopeptides,[159,160] b-turn dipeptide cores,[161,162] and β-aminoalcohols.[163,164]

VIII. Anti-IPMC Antibodies and Uses Therefor

Another aspect of the invention pertains to an antibody specifically reactive with a mammalian IPMC protein, e.g., a wild-type or mutated IPMC protein. For example, by using immunogens derived from an IPMC protein, e.g., based on the cDNA sequences, anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols.[165] A mammal, such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide (e.g., a mammalian IPMC polypeptide or an antigenic fragment which is capable of eliciting an antibody response, or a fusion protein as described above). Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other

techniques well known in the art. An immunogenic portion of an IPMC protein can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies. In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of an IPMC protein of a mammal, e.g., antigenic determinants of a protein set forth in SEQ ID NO:4 or 6 or closely related homologs (e.g., at least 90% homologous, and more preferably at least 94% homologous).

Following immunization of an animal with an antigenic preparation of an IPMC polypeptide, anti-IPMC antisera can be obtained and, if desired, polyclonal anti-IPMC antibodies isolated from the serum. To produce monoclonal antibodies, antibody-producing cells (lymphocytes) can be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique,[166] the human B cell hybridoma technique,[167] and the EBV-hybridoma technique to produce human monoclonal antibodies.[168] Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with a mammalian IPM polypeptide of the present invention and monoclonal antibodies isolated from a culture comprising such hybridoma cells. In one embodiment anti-human IPMC antibodies specifically react with the protein encoded by a nucleic acid having SEQ ID NO:4 or 6.

The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with one of the subject mammalian IPMC polypeptides. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, $F(ab)_2$ fragments can be generated by treating antibody with pepsin. The resulting $F(ab)_2$ fragment can be treated to reduce disulfide bridges to produce Fab fragments. The antibody of the present invention is further intended to include bispecific, single-chain, and chimeric and humanized molecules having affinity for an IPMC protein conferred by at least one CDR region of the antibody. In preferred embodiments, the antibody further comprises a label attached thereto and able to be

detected, (e.g., the label can be a radioisotope, fluorescent compound, enzyme or enzyme co-factor).

Anti-IPMC antibodies can be used, e.g., to monitor IPMC protein levels in an individual for determining, e.g., whether a subject has a disease or condition associated with an aberrant IPMC protein level, or allowing determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. The level of IPM polypeptides may be measured from cells in bodily fluid, such as in blood samples.

Another application of anti-IPMC antibodies of the present invention is in the immunological screening of cDNA libraries constructed in expression vectors such as λgt11, λgt18-23, λZAP, and λORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, λgt11 can produce fusion proteins whose amino termini consist of ß-galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of an IPMC protein, e.g., other orthologs of a particular IPMC protein or other paralogs from the same species, can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with anti-IPMC antibodies. Positive phage detected by this assay can then be isolated from the infected plate. Thus, the presence of IPMC protein homologs can be detected and cloned from other animals, as can alternate isoforms (including splice variants) from humans.

In another aspect, the invention provides methods for identifying autoantibodies to IPMC components. For example, naturally occurring autoantibodies may be caused by an autoimmune disease involving antibodies directed at IPMC proteins or nucleic acids. The IPMC nucleic acids and proteins disclosed herein provide assays (e.g., immunoassays) for the detection, isolation and characterization of anti-IPMC antibodies. Characterization of IPMC autantibodies encompasses the characterization and isolation of the IPMC autoantibody antigen or epitope.

IX. Transgenic animals

The invention further provides for transgenic animals, which can be used for a variety of purposes, e.g., to identify IPMC therapeutics. Transgenic animals of the invention include non-human animals containing a heterologous IPMC gene or fragment thereof under the

control of an IPMC promoter or under the control of a heterologous promoter. Accordingly, the transgenic animals of the invention can be animals expressing a transgene encoding a wild-type IPMC protein or fragment thereof or variants thereof, including mutants and polymorphic variants thereof. Such animals can be used, e.g., to determine the effect of a difference in amino acid sequence of an IPMC protein from the sequence set forth in SEQ ID NO:4 or 6, such as a polymorphic difference. These animals can also be used to determine the effect of expression of an IPMC protein in a specific site or for identifying IPMC therapeutics or confirming their activity *in vivo*.

The transgenic animals can also be animals containing a transgene, such as reporter gene, under the control of an IPMC promoter or fragment thereof. These animals are useful, e.g., for identifying compound that modulate production of IPM, such as by modulating IPMC gene expression. An IPMC gene promoter can be isolated, e.g., by screening of a genomic library with an IPMC cDNA fragment and characterized according to methods known in the art. In a preferred embodiment of the present invention, the transgenic animal containing said IPMC reporter gene is used to screen a class of bioactive molecules known for their ability to modulate IPMC protein expression. Yet other non-human animals within the scope of the invention include those in which the expression of the endogenous IPMC gene has been mutated or "knocked out". A "knock out" animal is one carrying a homozygous or heterozygous deletion of a particular gene or genes. These animals could be useful to determine whether the absence of an IPMC protein results in a specific phenotype, in particular whether these mice have or are likely to develop a specific disease, such as high susceptibility to macular degeneration. Furthermore these animals are useful in screens for drugs which alleviate or attenuate the disease condition resulting from the mutation of an IPMC gene as outlined below. These animals are also useful for determining the effect of a specific amino acid difference, or allelic variation, in an IPMC gene. That is, the IPMC knock out animals can be crossed with transgenic animals expressing, e.g., a mutated form or allelic variant of IPM, thus resulting in an animal which expresses only the mutated protein and not the wild-type IPMC protein.

In a preferred embodiment of this aspect of the invention, a transgenic IPMC knock-out mouse, carrying the mutated IPMC locus on one or both of its chromosomes, is used as a

model system for transgenic or drug treatment of the condition resulting from loss of IPMC expression.

Methods for obtaining transgenic and knockout non-human animals are well known in the art. Knock out mice are generated by homologous integration of a "knock out" construct into a mouse embryonic stem cell chromosome which encodes the gene to be knocked out. In one embodiment, gene targeting, which is a method of using homologous recombination to modify an animal's genome, can be used to introduce changes into cultured embryonic stem cells (ES cells). By targeting a IPMC gene of interest in ES cells, these changes can be introduced into the germlines of animals to generate chimeras. The gene targeting procedure is accomplished by introducing into tissue culture cells a DNA targeting construct that includes a segment homologous to a target IPMC locus, and which also includes an intended sequence modification to the IPMC genomic sequence (e.g., insertion, deletion, point mutation). The treated cells are then screened for accurate targeting to identify and isolate those which have been properly targeted.

Gene targeting in ES cells is in fact a scheme contemplated by the present invention as a means for disrupting a IPMC gene function through the use of a targeting transgene construct designed to undergo homologous recombination with one or more IPMC genomic sequences. The targeting construct can be arranged so that, upon recombination with an element of a IPMC gene, a positive selection marker is inserted into (or replaces) coding sequences of the gene. The inserted sequence functionally disrupts the IPMC gene, while also providing a positive selection trait. Exemplary IPMC targeting constructs are described in more detail below.

Generally, the ES cells used to produce the knockout animals are of the same species as the knockout animal to be generated. Thus for example, mouse embryonic stem cells can usually be used for generation of knockout mice. Embryonic stem cells are generated and maintained using methods well known to the skilled artisan.[169] Any line of ES cells can be used, however, the line chosen is typically selected for the ability of the cells to integrate into and become part of the germ line of a developing embryo so as to create germ line transmission of the knockout construct. Thus, any ES cell line that is believed to have this capability is suitable for use herein. One mouse strain that is typically used for production of

ES cells, is the 129J strain. Another ES cell line is murine cell line D3.[170] Still another preferred ES cell line is the WW6 cell line.[171] The cells are cultured and prepared for knockout construct insertion using methods well known to the skilled artisan.[172,173,174]

A knock out construct refers to a uniquely configured fragment of nucleic acid which is introduced into a stem cell line and allowed to recombine with the genome at the chromosomal locus of the gene of interest to be mutated. Thus a given knock out construct is specific for a given gene to be targeted for disruption. Nonetheless, many common elements exist among these constructs and these elements are well known in the art. A typical knock out construct contains nucleic acid fragments of not less than about 0.5 kb nor more than about 10.0 kb from both the 5' and the 3' ends of the genomic locus which encodes the gene to be mutated. These two fragments are separated by an intervening fragment of nucleic acid which encodes a positive selectable marker, such as the neomycin resistance gene (neo^R). The resulting nucleic acid fragment, consisting of a nucleic acid from the extreme 5' end of the genomic locus linked to a nucleic acid encoding a positive selectable marker which is in turn linked to a nucleic acid from the extreme 3' end of the genomic locus of interest, omits most of the coding sequence for the IPMC gene or other gene of interest to be knocked out. When the resulting construct recombines homologously with the chromosome at this locus, it results in the loss of the omitted coding sequence, otherwise known as the structural gene, from the genomic locus. A stem cell in which such a rare homologous recombination event has taken place can be selected for by virtue of the stable integration into the genome of the nucleic acid of the gene encoding the positive selectable marker and subsequent selection for cells expressing this marker gene in the presence of an appropriate drug (neomycin in this example).

Variations on this basic technique also exist and are well known in the art. For example, a "knock-in" construct refers to the same basic arrangement of a nucleic acid encoding a 5' genomic locus fragment linked to nucleic acid encoding a positive selectable marker which in turn is linked to a nucleic acid encoding a 3' genomic locus fragment, but which differs in that none of the coding sequence is omitted and thus the 5' and the 3' genomic fragments used were initially contiguous before being disrupted by the introduction of the nucleic acid encoding the positive selectable marker gene. This "knock-in" type of construct is thus very useful for the construction of mutant transgenic animals when only a limited

region of the genomic locus of the gene to be mutated, such as a single exon, is available for cloning and genetic manipulation. Alternatively, the "knock-in" construct can be used to specifically eliminate a single functional domain of the targetted gene, resulting in a transgenic animal which expresses a polypeptide of the targetted gene which is defective in one function, while retaining the function of other domains of the encoded polypeptide. This type of "knock-in" mutant frequently has the characteristic of a so-called "dominant negative" mutant because, especially in the case of proteins which homomultimerize, it can specifically block the action of (or "poison") the polypeptide product of the wild-type gene from which it was derived. In a variation of the knock-in technique, a marker gene is integrated at the genomic locus of interest such that expression of the marker gene comes under the control of the transcriptional regulatory elements of the targeted gene. A marker gene is one that encodes an enzyme whose activity can be detected (e.g., β-galactosidase), the enzyme substrate can be added to the cells under suitable conditions, and the enzymatic activity can be analyzed. One skilled in the art is familiar with other useful markers and the means for detecting their presence in a given cell. All such markers are contemplated as being included within the scope of the teaching of this invention.

As mentioned above, the homologous recombination of the above described "knock out" and "knock in" constructs is very rare and frequently such a construct inserts nonhomologously into a random region of the genome where it has no effect on the gene which has been targeted for deletion, and where it can potentially recombine so as to disrupt another gene which was otherwise not intended to be altered. Such nonhomologous recombination events can be selected against by modifying the above mentioned knock out and knock in constructs so that they are flanked by negative selectable markers at either end (particularly through the use of two allelic variants of the thymidine kinase gene, the polypeptide product of which can be selected against in expressing cell lines in an appropriate tissue culture medium well known in the art - i.e. one containing a drug such as 5-bromodeoxyuridine). Thus a preferred embodiment of such a knock out or knock in construct of the invention consist of a nucleic acid encoding a negative selectable marker linked to a nucleic acid encoding a 5' end of a genomic locus linked to a nucleic acid of a positive selectable marker which in turn is linked to a nucleic acid encoding a 3' end of the same

genomic locus which in turn is linked to a second nucleic acid encoding a negative selectable marker. Nonhomologous recombination between the resulting knock out construct and the genome usually result in the stable integration of one or both of these negative selectable marker genes and hence cells which have undergone nonhomologous recombination can be selected against by growth in the appropriate selective media (e.g., media containing a drug such as 5-bromodeoxyuridine for example). Simultaneous selection for the positive selectable marker and against the negative selectable marker results in a vast enrichment for clones in which the knock out construct has recombined homologously at the locus of the gene intended to be mutated. The presence of the predicted chromosomal alteration at the targeted gene locus in the resulting knock out stem cell line can be confirmed by means of Southern blot analytical techniques which are well known to those familiar in the art. Alternatively, PCR can be used.

Each knockout construct to be inserted into the cell must first be in the linear form. Therefore, if the knockout construct has been inserted into a vector (described *infra*), linearization is accomplished by digesting the DNA with a suitable restriction endonuclease selected to cut only within the vector sequence and not within the knockout construct sequence.

For insertion, the knockout construct is added to the ES cells under appropriate conditions for the insertion method chosen, as is known to the skilled artisan. For example, if the ES cells are to be electroporated, the ES cells and knockout construct DNA are exposed to an electric pulse using an electroporation machine and following the manufacturer's guidelines for use. After electroporation, the ES cells are typically allowed to recover under suitable incubation conditions. The cells are then screened for the presence of the knock out construct as explained above. Where more than one construct is to be introduced into the ES cell, each knockout construct can be introduced simultaneously or one at a time.

After suitable ES cells containing the knockout construct in the proper location have been identified by the selection techniques outlined above, the cells can be inserted into an embryo. Insertion may be accomplished in a variety of ways known to the skilled artisan, however a preferred method is by microinjection. For microinjection, about 10-30 cells are collected into a micropipet and injected into embryos that are at the proper stage of

development to permit integration of the foreign ES cell containing the knockout construct into the developing embryo. For instance, the transformed ES cells can be microinjected into blastocytes. The suitable stage of development for the embryo used for insertion of ES cells is very species dependent, however for mice it is about 3.5 days. The embryos are obtained by perfusing the uterus of pregnant females. Suitable methods for accomplishing this are known to the skilled artisan.[173]

While any embryo of the right stage of development is suitable for use, preferred embryos are male. In mice, the preferred embryos also have genes coding for a coat color that is different from the coat color encoded by the ES cell genes. In this way, the offspring can be screened easily for the presence of the knockout construct by looking for mosaic coat color (indicating that the ES cell was incorporated into the developing embryo). Thus, for example, if the ES cell line carries the genes for white fur, the embryo selected carries the genes for black or brown fur.

After the ES cell has been introduced into the embryo, the embryo may be implanted into the uterus of a pseudopregnant foster mother for gestation. While any foster mother may be used, the foster mother is typically selected for her ability to breed and reproduce well, and for her ability to care for the young. Such foster mothers are typically prepared by mating with vasectomized males of the same species. The stage of the pseudopregnant foster mother is important for successful implantation, and it is species dependent. For mice, this stage is about 2-3 days pseudopregnant.

Offspring that are born to the foster mother may be screened initially for mosaic coat color where the coat color selection strategy (as described above, and in the appended examples) has been employed. In addition, or as an alternative, DNA from tail tissue of the offspring may be screened for the presence of the knockout construct using Southern blots and/or PCR as described above. Offspring that appear to be mosaics may then be crossed to each other, if they are believed to carry the knockout construct in their germ line, in order to generate homozygous knockout animals. Homozygotes may be identified by Southern blotting of equivalent amounts of genomic DNA from mice that are the product of this cross, as well as mice that are known heterozygotes and wild type mice.

Other means of identifying and characterizing the knockout offspring are available. For example, Northern blots can be used to probe the mRNA for the presence or absence of transcripts encoding either the gene knocked out, the marker gene, or both. In addition, Western blots can be used to assess the level of expression of the IPMC gene knocked out in various tissues of the offspring by probing the Western blot with an antibody against the particular IPMC protein, or an antibody against the marker gene product, where this gene is expressed. Finally, *in situ* analysis (such as fixing the cells and labeling with antibody) and/or FACS (fluorescence activated cell sorting) analysis of various cells from the offspring can be conducted using suitable antibodies to look for the presence or absence of the knockout construct gene product.

Yet other methods of making knock-out or disruption transgenic animals are also generally known.[174] Recombinase dependent knockouts can also be generated, e.g. by homologous recombination to insert target sequences, such that tissue specific and/or temporal control of inactivation of an IPMC gene can be controlled by recombinase sequences (described *infra*).

Animals containing more than one knockout construct and/or more than one transgene expression construct are prepared in any of several ways. The preferred manner of preparation is to generate a series of mammals, each containing one of the desired transgenic phenotypes. Such animals are bred together through a series of crosses, backcrosses and selections, to ultimately generate a single animal containing all desired knockout constructs and/or expression constructs, where the animal is otherwise congenic (genetically identical) to the wild type except for the presence of the knockout construct(s) and/or transgene(s) .

A IPMC transgene can encode the wild-type form of the protein, or can encode homologs thereof, including both agonists and antagonists, as well as antisense constructs. In preferred embodiments, the expression of the transgene is restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, *cis*-acting sequences that control expression in the desired pattern. In the present invention, such mosaic expression of a IPMC protein can be essential for many forms of lineage analysis and can additionally provide a means to assess the effects of, for example, lack of IPMC protein expression which might grossly alter the structure and integrity of retinal tissue. Toward this end, tissue-specific

regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory sequences.

Genetic techniques, which allow for the expression of transgenes can be regulated via site-specific genetic manipulation *in vivo,* are known to those skilled in the art. For instance, genetic systems are available which allow for the regulated expression of a recombinase that catalyzes the genetic recombination of a target sequence. As used herein, the phrase "target sequence" refers to a nucleotide sequence that is genetically recombined by a recombinase. The target sequence is flanked by recombinase recognition sequences and is generally either excised or inverted in cells expressing recombinase activity. Recombinase catalyzed recombination events can be designed such that recombination of the target sequence results in either the activation or repression of expression of one of the subject IPMC proteins. For example, excision of a target sequence which interferes with the expression of a recombinant IPMC gene, such as one which encodes an antagonistic homolog or an antisense transcript, can be designed to activate expression of that gene. This interference with expression of the protein can result from a variety of mechanisms, such as spatial separation of the IPMC gene from the promoter element or an internal stop codon. Moreover, the transgene can be made wherein the coding sequence of the gene is flanked by recombinase recognition sequences and is initially transfected into cells in a 3' to 5' orientation with respect to the promoter element. In such an instance, inversion of the target sequence reorients the subject gene by placing the 5' end of the coding sequence in an orientation with respect to the promoter element which allow for promoter driven transcriptional activation.

The transgenic animals of the present invention all include within a plurality of their cells a transgene of the present invention, which transgene alters the phenotype of the "host cell" with respect to regulation of cell growth, death and/or differentiation. Since it is possible to produce transgenic organisms of the invention utilizing one or more of the transgene constructs described herein, a general description is given of the production of transgenic organisms by referring generally to exogenous genetic material. This general description can

be adapted by those skilled in the art in order to incorporate specific transgene sequences into organisms utilizing the methods and materials described below.

In an illustrative embodiment, either the *cre/loxP* recombinase system of bacteriophage P1,[175,176] or the FLP recombinase system of *Saccharomyces cerevisiae*[177,178] can be used to generate *in vivo* site-specific genetic recombination systems. Cre recombinase catalyzes the site-specific recombination of an intervening target sequence located between *loxP* sequences. *loxP* sequences are 34 base pair nucleotide repeat sequences to which the Cre recombinase binds and are required for Cre recombinase mediated genetic recombination. The orientation of *loxP* sequences determines whether the intervening target sequence is excised or inverted when Cre recombinase is present;[179] catalyzing the excision of the target sequence when the *loxP* sequences are oriented as direct repeats and catalyzes inversion of the target sequence when *loxP* sequences are oriented as inverted repeats.

Accordingly, genetic recombination of the target sequence is dependent on expression of the Cre recombinase. Expression of the recombinase can be regulated by promoter elements which are subject to regulatory control, e.g., tissue-specific, developmental stage-specific, inducible or repressible by externally added agents. This regulated control results in genetic recombination of the target sequence only in cells where recombinase expression is mediated by the promoter element. Thus, the activation of expression of a recombinant IPMC protein can be regulated via control of recombinase expression.

Use of the *cre/loxP* recombinase system to regulate expression of a recombinant IPMC protein requires the construction of a transgenic animal containing transgenes encoding both the Cre recombinase and the subject protein. Animals containing both the Cre recombinase and a recombinant IPMC gene can be provided through the construction of "double" transgenic animals. A convenient method for providing such animals is to mate two transgenic animals each containing a transgene, e.g., an IPMC gene and recombinase gene.

One advantage derived from initially constructing transgenic animals containing an IPMC transgene in a recombinase-mediated expressible format derives from the likelihood that the subject protein, whether agonistic or antagonistic, can be deleterious upon expression in the transgenic animal. In such an instance, a founder population, in which the subject transgene is silent in all tissues, can be propagated and maintained. Individuals of this founder

population can be crossed with animals expressing the recombinase in, for example, one or more tissues and/or a desired temporal pattern. Thus, the creation of a founder population in which, for example, an antagonistic IPMC transgene is silent allows the study of progeny from that founder in which disruption of IPMC protein mediated induction in a particular tissue or at certain developmental stages would result in, for example, a lethal phenotype.

Similar conditional transgenes can be provided using prokaryotic promoter sequences which require prokaryotic proteins to be simultaneous expressed in order to facilitate expression of the IPMC transgene. Exemplary promoters and the corresponding trans-activating prokaryotic proteins are given in U.S. Patent No. 4,833,080.

Moreover, expression of the conditional transgenes can be induced by gene therapy-like methods wherein a gene encoding the trans-activating protein, e.g. a recombinase or a prokaryotic protein, is delivered to the tissue and caused to be expressed, such as in a cell-type specific manner. By this method, a IPMC transgene could remain silent into adulthood until "turned on" by the introduction of the trans-activator.

In an exemplary embodiment, the "transgenic non-human animals" of the invention are produced by introducing transgenes into the germline of the non-human animal. Embryonal target cells at various developmental stages can be used to introduce transgenes. Different methods are used depending on the stage of development of the embryonal target cell. The specific line(s) of any animal used to practice this invention are selected for general good health, good embryo yields, good pronuclear visibility in the embryo, and good reproductive fitness. In addition, the haplotype is a significant factor. For example, when transgenic mice are to be produced, strains such as C57BL/6 or FVB lines are often used (Jackson Laboratory, Bar Harbor, ME). Preferred strains are those with H-2^b, H-2^d or H-2^q haplotypes such as C57BL/6 or DBA/1. The line(s) used to practice this invention may themselves be transgenics, and/or may be knockouts (i.e., obtained from animals which have one or more genes partially or completely suppressed).

In one embodiment, the transgene construct is introduced into a single stage embryo. The zygote is the best target for micro-injection. In the mouse, the male pronucleus reaches the size of approximately 20 micrometers in diameter which allows reproducible injection of 1-2pl of DNA solution. The use of zygotes as a target for gene transfer has a major advantage

in that in most cases the injected DNA is incorporated into the host gene before the first cleavage.[180] As a consequence, all cells of the transgenic animal carries the incorporated transgene. This in general is also reflected in the efficient transmission of the transgene to offspring of the founder since 50% of the germ cells are harbored in the transgene.

Normally, fertilized embryos are incubated in suitable media until the pronuclei appear. At about this time, the nucleotide sequence comprising the transgene is introduced into the female or male pronucleus as described below. In some species such as mice, the male pronucleus is preferred. It is most preferred that the exogenous genetic material be added to the male DNA complement of the zygote prior to its being processed by the ovum nucleus or the zygote female pronucleus. It is thought that the ovum nucleus or female pronucleus release molecules which affect the male DNA complement, perhaps by replacing the protamines of the male DNA with histones, thereby facilitating the combination of the female and male DNA complements to form the diploid zygote.

Thus, it is preferred that the exogenous genetic material be added to the male complement of DNA or any other complement of DNA prior to its being affected by the female pronucleus. For example, the exogenous genetic material is added to the early male pronucleus, as soon as possible after the formation of the male pronucleus, which is when the male and female pronuclei are well separated and both are located close to the cell membrane. Alternatively, the exogenous genetic material could be added to the nucleus of the sperm after it has been induced to undergo decondensation. Sperm containing the exogenous genetic material can then be added to the ovum or the decondensed sperm could be added to the ovum with the transgene constructs being added as soon as possible thereafter.

Introduction of the transgene nucleotide sequence into the embryo may be accomplished by any means known in the art such as, for example, microinjection, electroporation, or lipofection. Following introduction of the transgene nucleotide sequence into the embryo, the embryo may be incubated *in vitro* for varying amounts of time, or reimplanted into the surrogate host, or both. *In vitro* incubation to maturity is within the scope of this invention. One common method is to incubate the embryos *in vitro* for about 1-7 days, depending on the species, and then reimplant them into the surrogate host.

For the purposes of this invention a zygote is essentially the formation of a diploid cell which is capable of developing into a complete organism. Generally, the zygote can be comprised of an egg containing a nucleus formed, either naturally or artificially, by the fusion of two haploid nuclei from a gamete or gametes. Thus, the gamete nuclei must be ones which are naturally compatible, i.e., ones which result in a viable zygote capable of undergoing differentiation and developing into a functioning organism. Generally, a euploid zygote is preferred. If an aneuploid zygote is obtained, then the number of chromosomes should not vary by more than one with respect to the euploid number of the organism from which either gamete originated.

In addition to similar biological considerations, physical ones also govern the amount (e.g., volume) of exogenous genetic material which can be added to the nucleus of the zygote or to the genetic material which forms a part of the zygote nucleus. If no genetic material is removed, then the amount of exogenous genetic material which can be added is limited by the amount which are absorbed without being physically disruptive. Generally, the volume of exogenous genetic material inserted do not exceed about 10 picoliters. The physical effects of addition must not be so great as to physically destroy the viability of the zygote. The biological limit of the number and variety of DNA sequences vary depending upon the particular zygote and functions of the exogenous genetic material and are readily apparent to one skilled in the art, because the genetic material, including the exogenous genetic material, of the resulting zygote must be biologically capable of initiating and maintaining the differentiation and development of the zygote into a functional organism.

The number of copies of the transgene constructs which are added to the zygote is dependent upon the total amount of exogenous genetic material added and is the amount which enables the genetic transformation to occur. Theoretically only one copy is required; however, generally, numerous copies are utilized, for example, 1,000-20,000 copies of the transgene construct, in order to insure that one copy is functional. There is often an advantage to having more than one functioning copy of each of the inserted exogenous DNA sequences to enhance the phenotypic expression of the exogenous DNA sequences.

Any technique which allows for the addition of the exogenous genetic material into nucleic genetic material can be utilized so long as it is not destructive to the cell, nuclear

membrane or other existing cellular or genetic structures. The exogenous genetic material is preferentially inserted into the nucleic genetic material by microinjection. Microinjection of cells and cellular structures is known and is used in the art.

Reimplantation is accomplished using standard methods. Usually, the surrogate host is anesthetized, and the embryos are inserted into the oviduct. The number of embryos implanted into a particular host varies according to species, but usually is comparable to the number of offspring the species naturally produces.

Transgenic offspring of the surrogate host may be screened for the presence and/or expression of the transgene by any suitable method. Screening is often accomplished by Southern blot or Northern blot analysis, using a probe that is complementary to at least a portion of the transgene. Western blot analysis using an antibody against the protein encoded by the transgene may be employed as an alternative or additional method for screening for the presence of the transgene product. Typically, DNA is prepared from tail tissue and analyzed by Southern analysis or PCR for the transgene. Alternatively, the tissues or cells believed to express the transgene at the highest levels are tested for the presence and expression of the transgene using Southern analysis or PCR, although any tissues or cell types may be used for this analysis.

Alternative or additional methods for evaluating the presence of the transgene include, without limitation, suitable biochemical assays such as enzyme and/or immunological assays, histological stains for particular marker or enzyme activities, flow cytometric analysis, and the like. Analysis of the blood may also be useful to detect the presence of the transgene product in the blood, as well as to evaluate the effect of the transgene on the levels of various types of blood cells and other blood constituents.

Progeny of the transgenic animals may be obtained by mating the transgenic animal with a suitable partner, or by *in vitro* fertilization of eggs and/or sperm obtained from the transgenic animal. Where mating with a partner is to be performed, the partner may or may not be transgenic and/or a knockout; where it is transgenic, it may contain the same or a different transgene, or both. Alternatively, the partner may be a parental line. Where *in vitro* fertilization is used, the fertilized embryo may be implanted into a surrogate host or incubated

in vitro, or both. Using either method, the progeny may be evaluated for the presence of the transgene using methods described above, or other appropriate methods.

The transgenic animals produced in accordance with the present invention includes exogenous genetic material. As set out above, the exogenous genetic material will, in certain embodiments, be a DNA sequence which results in the production of a IPMC protein (either agonistic or antagonistic), and antisense transcript, or a IPMC mutant. Further, in such embodiments the sequence are attached to a transcriptional control element, e.g., a promoter, which preferably allows the expression of the transgene product in a specific type of cell.

Retroviral infection can also be used to introduce transgene into a non-human animal. The developing non-human embryo can be cultured *in vitro* to the blastocyst stage. During this time, the blastomeres can be targets for retroviral infection.[181] Efficient infection of the blastomeres is obtained by enzymatic treatment to remove the zona pellucida.[174] The viral vector system used to introduce the transgene is typically a replication-defective retrovirus carrying the transgene.[182,183] Transfection is easily and efficiently obtained by culturing the blastomeres on a monolayer of virus-producing cells.[183,184] Alternatively, infection can be performed at a later stage. Virus or virus-producing cells can be injected into the blastocoele.[185] Most of the founders are mosaic for the transgene since incorporation occurs only in a subset of the cells which formed the transgenic non-human animal. Further, the founder may contain various retroviral insertions of the transgene at different positions in the genome which generally segregate in the offspring. In addition, it is also possible to introduce transgenes into the germ line by intrauterine retroviral infection of the midgestation embryo.[185]

A third type of target cell for transgene introduction is the embryonal stem cell (ES). ES cells are obtained from pre-implantation embryos cultured *in vitro* and fused with embryos.[186,187,188,189] Transgenes can be efficiently introduced into the ES cells by DNA transfection or by retrovirus-mediated transduction. Such transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The ES cells thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal.[190]

X. Screening Assays for IPMC Therapeutics

The invention further provides screening methods for identifying IPMC therapeutics, e.g., for treating and/or preventing the development of a disease or condition associated with abnormal IPMC activity such as, e.g., a macular degenerative diseases.

An IPMC therapeutic can be any type of compound, including a native or synthetic protein, peptide, peptidomimetic, small molecule, nucleic acid, carbohydrate or lipid. Proteoglycan-associated oligosaccharides, functioning alone or in concert with their core proteins, mediate important biological events, including many that are crucial for the development, growth, function and/or survival of organisms. Further, IPMC degrading molecules may be identified using the nucleic acids and proteins of the invention. Such degrading molecules may be enzymes involved in the biosynthesis, addition, removal, or maintenance of post-translational modifications.

A nucleic acid can be, e.g., an IPMC gene, IPMC cDNA, an antisense nucleic acid, a ribozyme, or a triplex molecule. An IPMC therapeutic of the invention can be an agonist or an antagonist. Preferred IPMC agonists include IPMC genes or proteins or derivatives thereof which mimic at least one IPMC activity. Other preferred agonists include compounds which are capable of increasing the production of an IPMC protein in a cell, e.g., compounds capable of upregulating the expression of an IPMC gene, and compounds which are capable of enhancing an IPMC activity and/or the interaction of an IPMC protein with another molecule, such as a target peptide. Preferred IPMC antagonists include IPMC proteins which are dominant negative proteins. Other preferred antagonists include compounds which decrease or inhibit the production of an IPMC protein in a cell and compounds which are capable of downregulating expression of an IPMC gene, and compounds which are capable of downregulating an IPMC activity and/or interaction of an IPMC protein with another molecule.

The invention also provides screening methods for identifying IPMC agonist and antagonist compounds, comprising selecting compounds which are capable of interacting with an IPMC protein or with a molecule capable of interacting with an IPMC protein. In general, a molecule which is capable of interacting with an IPMC protein is referred to herein as an "IPMC binding partner."

The compounds of the invention can be identified using various assays depending on the type of compound and activity of the compound that is desired. In addition, as described herein, the test compounds can be further tested in animal models. Set forth below are at least some assays that can be used for identifying IPMC therapeutics. However, based on the instant disclosure, one of skill in the art could use additional assays for identifying IPMC therapeutics without requiring undue experimentation.

XI. Cell-free assays

Cell-free assays can be used to identify compounds which are capable of interacting with an IPMC protein or binding partner, to thereby modify the activity of the IPMC protein or binding partner. Such a compound can, e.g., modify the structure of an IPMC protein or binding partner and thereby effect its activity. Cell-free assays can also be used to identify compounds which modulate the interaction between an IPMC protein and an IPMC binding partner, such as a target peptide. In a preferred embodiment, cell-free assays for identifying such compounds consist essentially in a reaction mixture containing an IPMC protein and a test compound or a library of test compounds in the presence or absence of a binding partner. A test compound can be, e.g., a derivative of an IPMC binding partner, e.g., a biologically inactive target peptide, or a small molecule.

Accordingly, one exemplary screening assay of the present invention includes the steps of contacting an IPMC protein or functional fragment thereof or an IPMC binding partner with a test compound or library of test compounds and detecting the formation of complexes. For detection purposes, the molecule can be labeled with a specific marker and the test compound or library of test compounds labeled with a different marker. Interaction of a test compound with an IPMC protein or fragment thereof or IPMC binding partner can then be detected by determining the level of the two labels after an incubation step and a washing step. The presence of two labels after the washing step is indicative of an interaction.

An interaction between molecules can also be identified by using real-time BIA (Biomolecular Interaction Analysis, Pharmacia Biosensor AB) which detects surface plasmon resonance (SPR), an optical phenomenon. Detection depends on changes in the mass concentration of macromolecules at the biospecific interface, and does not require any labeling

of interactants. In one embodiment, a library of test compounds can be immobilized on a sensor surface, e.g., which forms one wall of a micro-flow cell. A solution containing the IPMC protein, functional fragment thereof, IPMC analog or IPMC binding partner is then flown continuously over the sensor surface. A change in the resonance angle as shown on a signal recording, indicates that an interaction has occurred. This technique is further described, e.g., in BIAtechnology Handbook by Pharmacia.

Another exemplary screening assay of the present invention includes the steps of (a) forming a reaction mixture including: (i) an IPMC protein, (ii) an IPMC binding partner, and (iii) a test compound; and (b) detecting interaction of the IPMC protein and the IPM binding protein. The IPMC protein and IPMC binding partner can be produced recombinantly, purified from a source, e.g., plasma, or chemically synthesized, as described herein. A statistically significant change (potentiation or inhibition) in the interaction of the IPMC protein and IPMC binding protein in the presence of the test compound, relative to the interaction in the absence of the test compound, indicates a potential agonist (mimetic or potentiator) or antagonist (inhibitor) of IPMC bioactivity for the test compound. The compounds of this assay can be contacted simultaneously. Alternatively, an IPMC protein can first be contacted with a test compound for an appropriate amount of time, following which the IPMC binding partner is added to the reaction mixture. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. In the control assay, isolated and purified IPMC protein or binding partner is added to a composition containing the IPMC binding partner or IPMC protein, and the formation of a complex is quantitated in the absence of the test compound.

Complex formation between an IPMC protein and an IPM binding partner may be detected by a variety of techniques. Modulation of the formation of complexes can be quantitated using, for example, detectably labeled proteins such as radiolabeled, fluorescently labeled, or enzymatically labeled IPMC proteins or IPMC binding partners, by immunoassay, or by chromatographic detection.

Typically, it is desirable to immobilize either IPMC protein or its binding partner to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as

well as to accommodate automation of the assay. Binding of IPMC protein to an IPMCbinding partner, can be accomplished in any vessel suitable for containing the reactants. Examples include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to a matrix. For example, glutathione-S-transferase/IPMC protein (GST/IPM) fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the IPMC binding partner, e.g. an ^{35}S-labeled IPMC binding partner, and the test compound, and the mixture incubated under conditions conducive to complex formation, e.g. at physiological conditions for salt and pH, though slightly more stringent conditions may be desired. Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly (e.g. beads placed in scintilant), or in the supernatant after the complexes are subsequently dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-PAGE, and the level of IPMC protein or IPMC binding partner found in the bead fraction quantitated from the gel using standard electrophoretic techniques such as described in the appended examples.

Other techniques for immobilizing proteins on matrices are also available for use in the subject assay. For instance, either IPMC protein or its cognate binding partner can be immobilized utilizing conjugation of biotin and streptavidin. For instance, biotinylated IPMC protein molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with IPMC protein can be derivatized to the wells of the plate, and IPMC protein trapped in the wells by antibody conjugation. As above, preparations of an IPMC binding protein and a test compound are incubated in the IPMC protein presenting wells of the plate, and the amount of complex trapped in the well can be quantitated. Exemplary methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the IPMC binding partner, or which are reactive with IPMC protein and compete with the binding partner; as well as enzyme-linked assays which rely on detecting an

enzymatic activity associated with the binding partner, either intrinsic or extrinsic activity. In the instance of the latter, the enzyme can be chemically conjugated or provided as a fusion protein with the IPMC binding partner. To illustrate, the IPMC binding partner can be chemically cross-linked or genetically fused with horseradish peroxidase, and the amount of polypeptide trapped in the complex can be assessed with a chromogenic substrate of the enzyme, e.g. 3,3'-diamino-benzadine tetrahydrochloride or 4-chloro-1-napthol. Likewise, a fusion protein comprising the polypeptide and glutathione-S-transferase can be provided, and complex formation quantitated by detecting the GST activity using 1-chloro-2,4-dinitrobenzene.[191]

For processes which rely on immunodetection for quantitating one of the proteins trapped in the complex, antibodies against the protein, such as anti-IPMC antibodies, can be used. Alternatively, the protein to be detected in the complex can be "epitope tagged" in the form of a fusion protein which includes, in addition to the IPMC sequence, a second polypeptide for which antibodies are readily available (e.g. from commercial sources). For instance, the GST fusion proteins described above can also be used for quantification of binding using antibodies against the GST moiety. Other useful epitope tags include myc-epitopes[84] which includes a 10-residue sequence from c-myc, as well as the pFLAG system (International Biotechnologies, Inc.) or the pEZZ-protein A system (Pharmacia, NJ).

Cell-free assays can also be used to identify compounds which interact with an IPMC protein and modulate an activity of an IPMC protein. Accordingly, in one embodiment, an IPMC protein is contacted with a test compound and the catalytic activity of IPMC protein is monitored. In one embodiment, the ability of the IPMC protein to bind a target molecule is determined. The binding affinity of the IPMC protein to a target molecule can be determined according to methods known in the art. Determination of the enzymatic activity of the IPMC protein can be performed with the aid of the substrate furanacryloyl-L-phenylalanyl-glycyl-glycine (FAPGG).[192,193]

XII. Cell based assays

In addition to cell-free assays, such as described above, IPMC proteins as provided by the present invention facilitate the generation of cell-based assays, e.g., for identifying small molecule agonists or antagonists. Cell based assays can be used, for example, to identify compounds which modulate expression of an IPMC gene, modulate translation of an IPMC mRNA, which modulate the stability of an IPMC mRNA or protein or which otherwise interfere with an interaction between an IPMC gene or protein and an IPMC binding partner. Accordingly, in one embodiment, a cell which is capable of producing IPMC protein is incubated with a test compound and the amount of IPMC protein produced in the cell medium is measured and compared to that produced from a cell which has not been contacted with the test compound. The specificity of the compound vis a vis IPMC protein can be confirmed by various control analysis, e.g., measuring the expression of one or more control genes. Compounds which can be tested include small molecules, proteins, and nucleic acids. In particular, this assay can be used to determine the efficacy of IPMC antisense molecules or ribozymes.

In another embodiment, the effect of a test compound on transcription of an IPMC gene is determined by transfection experiments using a reporter gene operatively linked to at least a portion of the promoter of an IPMC gene. A promoter region of a gene can be isolated, e.g., from a genomic library according to methods known in the art. The reporter gene can be any gene encoding a protein which is readily quantifiable, e.g., the luciferase or CAT gene. Such reporter gene are well known in the art.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

XIII. Predictive Medicine

The invention further features predictive medicines, which are based, at least in part, on the identity of the novel IPMC gene and alterations in the genes and related pathway genes, which affect the expression level and/or function of the encoded IPMC protein in a subject.

For example, as described herein, IPMC mutations that are particularly likely to cause or contribute to the development of diseases or conditions associated with abnormal IPMC

activity, such as retinal detachment or macular degeneration, are those mutations that negatively impact normal (wildtype) functioning of the IPM150 D domain that may be involved with the hyaluronan and EGF binding properties of IPMC protein. Examples of such mutations include: (i) upstream mutations that encode truncated transcripts that lack the D domain (e.g. a deletion encoding an IPMC protein transcript that is missing one or more of amino acid residues 735 to 743 of SEQ ID NO:4 and/or one or more of amino acid residues 688 to 731 of SEQ ID NO:4); (ii) missense mutations occurring within the D domain; or (iii) translocations that result in expression of only one copy of the IPMC protein (e.g. monosomy of chromosome 3), can result in a macular degeneration phenotype.

Information obtained using the diagnostic assays described herein (alone or in conjunction with information on another genetic defect, which contributes to the same disease) is useful for prognosing, diagnosing or confirming that a subject has a genetic defect (e.g. in an IPMC gene or in a gene that regulates the expression of an IPMC gene), which causes or contributes to the development of a disease or disorder such as retinal detachment or macular degeneration.

In addition, knowledge of the particular alteration or alterations, resulting in defective or deficient IPMC genes or proteins in an individual (the IPMC genetic profile), alone or in conjunction with information on other genetic defects contributing to a genetic disease or condition (e.g., the "macular degeneration genetic profile") allows customization of therapy to the individual's genetic profile, the goal of "pharmacogenomics". For example, an individual's IPMC genetic profile or the macular degeneration genetic profile, can enable a doctor to: (1) more effectively prescribe a drug that addresses the molecular basis of macular degeneration; and (2) better determine the appropriate dosage of a particular drug for the particular individual. For example, the expression level of IPMC proteins, alone or in conjunction with the expression level of other genes, known to contribute to the same disease, can be measured in many patients at various stages of the disease to generate a transcriptional or expression profile of the disease. Expression patterns of individual patients can then be compared to the expression profile of the disease to determine the appropriate drug and dose to administer to the patient.

The ability to target populations expected to show the highest clinical benefit, based on the IPMC protein or disease genetic profile, can enable: 1) the repositioning of marketed drugs with disappointing market results; 2) the rescue of drug candidates whose clinical development has been discontinued as a result of safety or efficacy limitations, which are patient subgroup-specific; and 3) an accelerated and less costly development for drug candidates and more optimal drug labeling (e.g. since the use of IPMC protein as a marker is useful for optimizing effective dose).

These and other methods are described in further detail in the following sections.

XIV. Prognostic and Diagnostic Assays

The present methods provide means for determining if a subject has (diagnostic) or is at risk of developing (prognostic) a disease, condition or disorder that is associated with an aberrant IPMC activity, e.g., an aberrant level of IPMC protein or an aberrant bioactivity, such as results in the development of macular degeneration.

Accordingly, the invention provides methods for determining whether a subject has or is likely to develop a disease or disorder associated with abnormal IPMC activity such as retinal detachment or macular degeneration, comprising determining the level of an IPMC gene or protein, an IPMC bioactivity and/or the presence of a mutation or particular polymorphic variant in the IPMC gene.

In one embodiment, the method comprises determining whether a subject has an abnormal mRNA and/or protein level of IPMC protein, such as by Northern blot analysis, reverse transcription-polymerase chain reaction (RT-PCR), *in situ* hybridization, immunoprecipitation, Western blot hybridization, or immunohistochemistry. According to the method, cells are obtained from a subject and the IPMC protein or mRNA level is determined and compared to the level of IPMC protein or mRNA level in a healthy subject. An abnormal level of IPMC polypeptide or mRNA level is likely to be indicative of an aberrant IPMC activity.

In another embodiment, the method comprises measuring at least one activity of IPMC protein as described herein. For example, maintenance of attachment of the RPE to the retina by an IPMC protein can be determined, e.g., as described herein. Comparison of the results

obtained with results from similar analysis performed on IPMC proteins from healthy subjects is indicative of whether a subject has an abnormal IPMC activity.

In preferred embodiments, the methods for determining whether a subject has or is at risk for developing a disease, which is caused by or contributed to by an aberrant IPMC activity is characterized as comprising detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of: (i) an alteration affecting the integrity of a gene encoding an IPMC polypeptide, or (ii) the mis-expression of the IPMC gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of: (i) a deletion of one or more nucleotides from an IPMC gene, (ii) an addition of one or more nucleotides to an IPMC gene, (iii) a substitution of one or more nucleotides of an IPMC gene, (iv) a gross chromosomal rearrangement of an IPMC gene, (v) a gross alteration in the level of a messenger RNA transcript of an IPMC gene, (vi) aberrant modification of an IPMC gene, such as of the methylation pattern of the genomic DNA, (vii) the presence of a non-wild type splicing pattern of a messenger RNA transcript of an IPMC gene, (viii) a non-wild type level of an IPM polypeptide, (ix) allelic loss of an IPMC gene, and/or (x) inappropriate post-translational modification of an IPM polypeptide. As set out below, the present invention provides a large number of assay techniques for detecting alterations in an IPMC gene. These methods include, but are not limited to, methods involving sequence analysis, Southern blot hybridization, restriction enzyme site mapping, and methods involving detection of the absence of nucleotide pairing between the nucleic acid to be analyzed and a probe. These and other methods are further described *infra*.

Specific diseases or disorders, e.g., genetic diseases or disorders, are associated with specific allelic variants of polymorphic regions of certain genes, which do not necessarily encode a mutated protein. Thus, the presence of a specific allelic variant of a polymorphic region of a gene, such as a single nucleotide polymorphism ("SNP"), in a subject can render the subject susceptible to developing a specific disease or disorder. Polymorphic regions in genes, e.g, IPMC genes, can be identified, by determining the nucleotide sequence of genes in populations of individuals. If a polymorphic region, e.g., SNP is identified, then the link with a specific disease can be determined by studying specific populations of individuals, e.g, individuals which developed a specific disease, such as retinal detachment or macular

degeneration. A polymorphic region can be located in any region of a gene, e.g., exons, in coding or non coding regions of exons, introns, and promoter region.

It is likely that IPMC genes comprise polymorphic regions, specific alleles of which may be associated with specific diseases or conditions or with an increased likelihood of developing such diseases or conditions. Thus, the invention provides methods for determining the identity of the allele or allelic variant of a polymorphic region of an IPMC gene in a subject, to thereby determine whether the subject has or is at risk of developing a disease or disorder that is associated with a specific allelic variant of a polymorphic region.

In an exemplary embodiment, there is provided a nucleic acid composition comprising a nucleic acid probe including a region of nucleotide sequence which is capable of hybridizing to a sense or antisense sequence of an IPMC gene or naturally occurring mutants thereof, or 5' or 3' flanking sequences naturally associated with the subject IPMC genes or naturally occurring mutants thereof. The nucleic acid of a cell is rendered accessible for hybridization, the probe is contacted with the nucleic acid of the sample, and the hybridization of the probe to the sample nucleic acid is detected. Such techniques can be used to detect alterations or allelic variants at either the genomic or mRNA level, including deletions, substitutions, etc., as well as to determine mRNA transcript levels.

A preferred detection method is allele specific hybridization using probes overlapping the mutation or polymorphic site and having about 5, 10, 20, 25, or 30 nucleotides around the mutation or polymorphic region. In a preferred embodiment of the invention, several probes capable of hybridizing specifically to allelic variants, such as single nucleotide polymorphisms, are attached to a solid phase support, e.g., a "chip". Oligonucleotides can be bound to a solid support by a variety of processes, including lithography. For example a chip can hold up to about 250,000 oligonucleotides. Mutation detection analysis using these chips comprising oligonucleotides, also termed "DNA probe arrays" is described.[194] In one embodiment, a chip comprises all the allelic variants of at least one polymorphic region of a gene. The solid phase support is then contacted with a test nucleic acid and hybridization to the specific probes is detected. Accordingly, the identity of numerous allelic variants of one or more genes can be identified in a simple hybridization experiment.

In certain embodiments, detection of the alteration comprises utilizing the probe/primer in a polymerase chain reaction (PCR),[195,196] such as anchor PCR or RACE PCR, or, alternatively, in a ligase chain reaction (LCR),[197,198] the latter of which can be particularly useful for detecting point mutations in an IPMC gene.[199] In a merely illustrative embodiment, the method includes the steps of (i) collecting a sample of cells from a patient, (ii) isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, (iii) contacting the nucleic acid sample with one or more primers which specifically hybridize to an IPMC gene under conditions such that hybridization and amplification of the IPMC gene (if present) occurs, and (iv) detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR, LCR or any other amplification procedure (e.g. self sustained sequence replication,[200] transcriptional amplification system,[201] or Q-Beta Replicase[202]), may be used as a preliminary step to increase the amount of sample on which any of the techniques for detecting mutations described herein can be performed.

In a preferred embodiment of the subject assay, mutations in, or allelic variants, of an IPMC gene from a sample cell are identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis. Moreover, the use of sequence specific ribozymes[203] can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence an IPMC gene and detect mutations by comparing the DNA sequence of the sample IPMC protein with the corresponding wild-type (control) DNA sequence. Exemplary sequencing reactions include those based on techniques developed by Maxim and Gilbert;[204] or Sanger.[205] It is also contemplated that any of a variety of automated sequencing procedures may be utilized when performing the subject assays,[206] including sequencing by mass spectrometry.[207,208,209] It is evident to one skilled in the art that, for certain embodiments, the occurrence of only one, two or three of the nucleic acid bases need be determined in the sequencing reaction. For instance, A-track or the like, e.g., where only one nucleic acid is detected, can be carried out.

In a further embodiment, protection from cleavage agents (such as a nuclease, hydroxylamine or osmium tetroxide and with piperidine) can be used to detect mismatched bases in RNA/RNA or RNA/DNA or DNA/DNA heteroduplexes.[210] In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes formed by hybridizing (labelled) RNA or DNA containing the wild-type IPMC nucleic acid sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which exists due to base pair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation.[211,212] In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in IPMC cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A·at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches.[213] According to an exemplary embodiment, a probe based on an IPMC DNA sequence, e.g., a wild-type IPMC DNA sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like.[214]

In other embodiments, alterations in electrophoretic mobility are used to identify mutations or the identity of the allelic variant of a polymorphic region in IPMC genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids.[215,216,217] Single-stranded DNA fragments of sample and control IPMC nucleic acids are denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to

sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labelled or detected with labelled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility.[218]

In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE).[219] When DGGE is used as the method of analysis, DNA are modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing agent gradient to identify differences in the mobility of control and sample DNA.[220]

Examples of other techniques for detecting point mutations or the identity of the allelic variant of a polymorphic region include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation or nucleotide difference (e.g., in allelic variants) is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found.[221,222] Such allele specific oligonucleotide hybridization techniques may be used to test one mutation or polymorphic region per reaction when oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations or polymorphic regions when the oligonucleotides are attached to the hybridizing membrane and hybridized with labelled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation or polymorphic region of interest in the center of the molecule[223] (so that amplification depends on differential hybridization) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension.[224] In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection.[225] It is

anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification.[226] In such cases, ligation occurs only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

In another embodiment, identification of the allelic variant is carried out using an oligonucleotide ligation assay (OLA).[197,227] The OLA protocol uses two oligonucleotides which are designed to be capable of hybridizing to abutting sequences of a single strand of a target. One of the oligonucleotides is linked to a separation marker, e.g,. biotinylated, and the other is detectably labeled. If the precise complementary sequence is found in a target molecule, the oligonucleotides hybridizes such that their termini abut, and create a ligation substrate. Ligation then permits the labeled oligonucleotide to be recovered using avidin, or another biotin ligand. Nickerson, D. A. et al.[228] have described a nucleic acid detection assay that combines attributes of PCR and OLA. In this method, PCR is used to achieve the exponential amplification of target DNA, which is then detected using OLA.

Several techniques based on this OLA method have been developed and can be used to detect specific allelic variants of a polymorphic region of an IPMC gene. For example, U.S. Patent No. 5,593,826 discloses an OLA using an oligonucleotide having 3'-amino group and a 5'-phosphorylated oligonucleotide to form a conjugate having a phosphoramidate linkage. In another variation of OLA described in Tobe et al.,[229] OLA combined with PCR permits typing of two alleles in a single microtiter well. By marking each of the allele-specific primers with a unique hapten, i.e. digoxigenin and fluorescein, each OLA reaction can be detected by using hapten specific antibodies that are labeled with different enzyme reporters, alkaline phosphatase or horseradish peroxidase. This system permits the detection of the two alleles using a high throughput format that leads to the production of two different colors.

The invention further provides methods for detecting single nucleotide polymorphisms in an IPMC gene. Because single nucleotide polymorphisms constitute sites of variation flanked by regions of invariant sequence, their analysis requires no more than the determination of the identity of the single nucleotide present at the site of variation and it is unnecessary to determine a complete gene sequence for each patient. Several methods have been developed to facilitate the analysis of such single nucleotide polymorphisms.

In one embodiment, the single base polymorphism can be detected by using a specialized exonuclease-resistant nucleotide, as disclosed, e.g., in Mundy, C. R.[230] According to the method, a primer complementary to the allelic sequence immediately 3' to the polymorphic site is permitted to hybridize to a target molecule obtained from a particular animal or human. If the polymorphic site on the target molecule contains a nucleotide that is complementary to the particular exonuclease-resistant nucleotide derivative present, then that derivative is incorporated onto the end of the hybridized primer. Such incorporation renders the primer resistant to exonuclease, and thereby permits its detection. Since the identity of the exonuclease-resistant derivative of the sample is known, a finding that the primer has become resistant to exonucleases reveals that the nucleotide present in the polymorphic site of the target molecule was complementary to that of the nucleotide derivative used in the reaction. This method has the advantage that it does not require the determination of large amounts of extraneous sequence data.

In another embodiment of the invention, a solution-based method is used for determining the identity of the nucleotide of a polymorphic site.[231] As in the Mundy[230] method, a primer is employed that is complementary to allelic sequences immediately 3' to a polymorphic site. The method determines the identity of the nucleotide of that site using labeled dideoxynucleotide derivatives, which, if complementary to the nucleotide of the polymorphic site becomes incorporated onto the terminus of the primer.

An alternative method, known as Genetic Bit Analysis or GBA™ is described by Goelet.[232] The method of Goelet[232] uses mixtures of labeled terminators and a primer that is complementary to the sequence 3' to a polymorphic site. The labeled terminator that is incorporated is thus determined by, and complementary to, the nucleotide present in the polymorphic site of the target molecule being evaluated. In contrast to the method of Cohen[231] the method of Goelet[232] is preferably a heterogeneous phase assay, in which the primer or the target molecule is immobilized to a solid phase.

Recently, several primer-guided nucleotide incorporation procedures for assaying polymorphic sites in DNA have been described.[233,234,235,236,237,238,239] These methods differ from GBA TM in that they all rely on the incorporation of labeled deoxynucleotides to discriminate between bases at a polymorphic site. In such a format, since the signal is proportional to the

number of deoxynucleotides incorporated, polymorphisms that occur in runs of the same nucleotide can result in signals that are proportional to the length of the run.[240]

For mutations that produce premature termination of protein translation, the protein truncation test (PTT) offers an efficient diagnostic approach.[241,242] For PTT, RNA is initially isolated from available tissue and reverse-transcribed, and the segment of interest is amplified by PCR. The products of reverse transcription PCR are then used as a template for nested PCR amplification with a primer that contains an RNA polymerase promoter and a sequence for initiating eukaryotic translation. After amplification of the region of interest, the unique motifs incorporated into the primer permit sequential *in vitro* transcription and translation of the PCR products. Upon sodium dodecyl sulfate-polyacrylamide gel electrophoresis of translation products, the appearance of truncated polypeptide signals the presence of a mutation that causes premature termination of translation. In a variation of this technique, DNA (as opposed to RNA) is used as a PCR template when the target region of interest is derived from a single exon.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid, primer set; and/or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving an IPMC protein.

Any cell type or tissue may be utilized in the diagnostics described below. In a preferred embodiment a bodily fluid, e.g., blood, is obtained from the subject to determine the presence of a mutation or the identity of the allelic variant of a polymorphic region of an IPMC gene. A bodily fluid, e.g, blood, can be obtained by known techniques (e.g. venipuncture). Alternatively, nucleic acid tests can be performed on dry samples (e.g. hair or skin). For prenatal diagnosis, fetal nucleic acid samples can be obtained from maternal blood as described in International Patent Application No. WO91/07660 to Bianchi. Alternatively, amniocytes or chorionic villi may be obtained for performing prenatal testing.

When using RNA or protein to determine the presence of a mutation or of a specific allelic variant of a polymorphic region of an IPMC gene, the cells or tissues that may be utilized must express the IPMC gene. Preferred cells for use in these methods include cardiac

cells (see Examples). Alternative cells or tissues that can be used, can be identified by determining the expression pattern of the specific IPMC gene in a subject, such as by Northern blot analysis.

Diagnostic procedures may also be performed *in situ* directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents may be used as probes and/or primers for such *in situ* procedures.[242]

In addition to methods which focus primarily on the detection of one nucleic acid sequence, profiles may also be assessed in such detection schemes. Fingerprint profiles may be generated, for example, by utilizing a differential display procedure, Northern analysis and/or RT-PCR.

Antibodies directed against wild type or mutant IPMC polypeptides or allelic variants thereof, which are discussed above, may also be used in disease diagnostics and prognostics. Such diagnostic methods, may be used to detect abnormalities in the level of IPMC polypeptide expression, or abnormalities in the structure and/or tissue, cellular, or subcellular location of an IPMC polypeptide. Structural differences may include, for example, differences in the size, electronegativity, or antigenicity of the mutant IPMC polypeptide relative to the normal IPMC polypeptide. Protein from the tissue or cell type to be analyzed may easily be detected or isolated using techniques which are well known to one of skill in the art, including but not limited to Western blot analysis.[129] The protein detection and isolation methods employed herein may also be such as those described in Harlow and Lane,[165] for example, which is incorporated herein by reference in its entirety.

This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody (see below) coupled with light microscopic, flow cytometric, or fluorimetric detection. The antibodies (or fragments thereof) useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for *in situ* detection of IPMC polypeptides. *In situ* detection may be accomplished by removing a histological specimen from a patient, and applying thereto a labeled antibody of the present invention. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of

such a procedure, it is possible to determine not only the presence of the IPMC polypeptide, but also its distribution in the examined tissue. Using the present invention, one of ordinary skill in the art can readily perceive that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such *in situ* detection.

Often a solid phase support or carrier is used as a support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art know of many other suitable carriers for binding antibody or antigen, or are able to ascertain the same by use of routine experimentation.

One means for labeling an anti-IPM polypeptide specific antibody is via linkage to an enzyme and use in an enzyme immunoassay (EIA).[243,244,245,246,247] The enzyme which is bound to the antibody reacts with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety which can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes which can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to

detect fingerprint gene wild type or mutant peptides through the use of a radioimmunoassay (RIA).[248] The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using fluorescence emitting metals such as ^{152}Eu, or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in, which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

Moreover, it is understood that any of the above methods for detecting alterations in a gene or gene product or polymorphic variants can be used to monitor the course of treatment or therapy.

XIV. Pharmacogenomics

Knowledge of the particular alteration or alterations, resulting in defective or deficient IPMC genes or proteins in an individual (an IPMC genetic profile), alone or in conjunction with information on other genetic defects contributing to the same disease (the genetic profile of the particular disease) allows a customization of the therapy for a particular disease to the individual's genetic profile, the goal of "pharmacogenomics". For example, subjects having a specific allele of an IPMC gene may or may not exhibit symptoms of a particular disease or be predisposed of developing symptoms of a particular disease. Further, if those subjects are symptomatic, they may or may not respond to a certain drug, e.g., a specific IPMC therapeutic, but may respond to another. Thus, generation of an IPMC genetic profile, (e.g., categorization of alterations in IPMC genes which are associated with the development of a disease or condition associated with abnormal IPMC activity such as retinal detachment or macular degeneration), from a population of subjects, who are symptomatic for a disease or condition that is caused by or contributed to by a defective and/or deficient IPMC gene and/or protein (an IPMC genetic population profile) and comparison of an individual's IPMC profile to the population profile, permits the selection or design of drugs that are expected to be safe and efficacious for a particular patient or patient population (i.e., a group of patients having the same genetic alteration).

For example, an IPMC population profile can be performed, by determining the IPMC profile, e.g., the identity of IPMC genes, in a patient population having a disease, which is caused by or contributed to by a defective or deficient IPMC gene. Optionally, the IPMC population profile can further include information relating to the response of the population to an IPMC therapeutic, using any of a variety of methods, including, monitoring: 1) the severity of symptoms associated with the IPMC related disease, 2) IPMC gene expression level, 3) IPMC mRNA level, and/or 4) IPMC protein level, and (iii) dividing or categorizing the population based on the particular genetic alteration or alterations present in its IPMC gene or an IPMC pathway gene. The IPMC genetic population profile can also, optionally, indicate those particular alterations in which the patient was either responsive or non-responsive to a particular therapeutic. This information or population profile, is then useful for predicting which individuals should respond to particular drugs, based on their individual IPMC profile.

In a preferred embodiment, the IPMC protein profile is a transcriptional or expression level profile and step (i) is comprised of determining the expression level of IPMC proteins, alone or in conjunction with the expression level of other genes, known to contribute to the same disease. The IPMC protein profile can be measured in many patients at various stages of the disease.

Pharmacogenomic studies can also be performed using transgenic animals. For example, one can produce transgenic mice, e.g., as described herein, which contain a specific allelic variant of an IPMC gene. These mice can be created, e.g, by replacing their wild-type IPMC gene with an allele of the human IPMC gene. The response of these mice to specific IPMC therapeutics can then be determined.

XVI. Monitoring of Effects of IPMC Therapeutics During Clinical Trials

The ability to target populations expected to show the highest clinical benefit, based on the IPMC protein or disease genetic profile, can enable: 1) the repositioning of marketed drugs with disappointing market results; 2) the rescue of drug candidates whose clinical development has been discontinued as a result of safety or efficacy limitations, which are patient subgroup-specific; and 3) an accelerated and less costly development for drug candidates and more optimal drug labeling (e.g. since the use of IPMC protein as a marker is useful for optimizing effective dose).

The treatment of an individual with an IPMC therapeutic can be monitored by determining IPMC characteristics, such as IPMC protein level or activity, IPMC mRNA level, and/or IPMC transcriptional level. Clinical tests useful for testing the efficacy of the treatment are well known in the art and include fundus fluorescein angiography (FFA), fundus photography (FP), electroretinogram (ERG), electrooculogram (EOG), visual fields, scanning laser opthalmoscopy (SLO), visual acuity measurements, dark adaptation measurements and/or any other standard ophthalmologic exam. These measurements indicates whether the treatment is effective or whether it should be adjusted or optimized. Thus, an IPMC protein can be used as a marker for the efficacy of a drug during clinical trials.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist,

peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a preadministration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of an IPMC protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the IPMC protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the IPMC protein, mRNA, or genomic DNA in the preadministration sample with the IPMC protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of the IPMC protein to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of IPMC protein to lower levels than detected, i.e., to decrease the effectiveness of the agent.

Cells of a subject may also be obtained before and after administration of an IPMC therapeutic to detect the level of expression of genes other than IPMC protein, to verify that the IPMC therapeutic does not increase or decrease the expression of genes which could be deleterious. This can be done, e.g., by using the method of transcriptional profiling. Thus, mRNA from cells exposed *in vivo* to an IPMC therapeutic and mRNA from the same type of cells that were not exposed to the IPMC therapeutic could be reverse transcribed and hybridized to a chip containing DNA from numerous genes, to thereby compare the expression of genes in cells treated and not treated with an IPMC therapeutic. If, for example, an IPMC therapeutic turns on the expression of a proto-oncogene in an individual, use of this particular IPMC therapeutic may be undesirable.

XVII. Kits

The invention further provides kits for use in diagnostics or prognostic methods for diseases or conditions associated with abnormal IPMC activity, such as retinal detachment, chorioretinal degeneration, retinal degeneration or macular degeneration or for determining which IPMC therapeutic should be administered to a subject, for example, by detecting the

presence of IPMC mRNA or protein in a biological sample. The kit may detect abnormal levels, form or activity of IPMC protein, RNA or a breakdown product of an IPMC protein or RNA. In an embodiment of the invention the kit detects autoantibodies specific for an IPMC protein, peptide or nucleic acid. For example, the kit can comprise a labeled compound or agent capable of detecting IPMC protein or mRNA in a biological sample; means for determining the amount of IPMC protein in the sample; and means for comparing the amount of IPMC protein in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect IPMC mRNA or protein. Such a kit can comprise, e.g., one or more nucleic acid probes capable of hybridizing specifically to at least a portion of an IPMC gene or allelic variant thereof, or mutated form thereof. Preferably the kit comprises at least one oligonucleotide primer capable of differentiating between a normal IPMC gene and an IPMC gene with one or more nucleotide differences.

XVIII. Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject having a disease or condition associated with abnormal IPMC bioactivity such as retinal detachment or macular degeneration. Subjects at risk for such a disease can be identified by a diagnostic or prognostic assay, e.g., as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the IPMC abnormality, e.g., such that development of the retinal detachment or macular degeneration is prevented or, alternatively, delayed in its progression. In general, the prophylactic or therapeutic methods comprise administering to the subject an effective amount of a compound which is capable of agonizing a wildtype IPMC activity or antagonizing a mutant (defective) IPMC activity. Examples of suitable compounds include the antagonists, agonists or homologs described in detail herein.

XIX. Effective Dose

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the Ld_{50} (the dose lethal to 50% of the population) and the Ed_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

XX. Formulation and Use

Pharmaceutical compositions for use in accordance with the present invention may be formulated in a conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by, for example, eye drops, injection, inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

For such therapy, the compounds of the invention can be formulated for a variety of modes of administration, including systemic and topical or localized administration. Techniques and formulations generally may be found in Remmington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. A preferred method of administration is an eye drop. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, and subcutaneous. For injection, the compounds of the invention can be formulated in liquid solutions, preferably in physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition, the compounds may be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included.

Preferred methods of administration include choroidal injection, transscleral injection or placing a scleral patch, and selective arterial catheterization. Other preferred deliveries are intraocular, including transretinal, subconjunctival bulbar, scleral pocket and scleral cutdown injections. The agent can be alternatively administered intravascularly, such as intravenously (IV) or intraarterially.

Techniques for choroidal injection and scleral patching are similar. The clinician uses a local approach to the eye after initiation of appropriate anesthesia, including painkillers and ophthalmoplegics. A needle containing the therapeutic compound is directed into the patient's choroid or sclera and inserted under sterile conditions. When the needle is properly positioned the compound is injected into either or both of the choroid or sclera. When using either of these methods, the clinician may choose a sustained release or longer acting formulation. Thus, the procedure may need repetition only every several months or several years, depending on the patient's tolerance of the treatment and response.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulfate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they

may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., ationd oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

The IPMC therapeutic may be administered alone or in combination with other molecules known to have a beneficial effect on retinal attachment or damaged retinal tissue, including molecules capable of tissue repair and regeneration and/or inhibiting inflammation. Examples of useful cofactors include basic fibroblast growth factor (bFGF),[249] ciliary neurotrophic factor (CNTF),[249] axokine (a mutein of CNTF),[249] leukemia inhibitory factor (LIF),[249] neutrotrophin 3 (NT-3),[249] neurotrophin-4 (NT-4),[249] nerve growth factor (NGF),[249] insulin-like growth factor II,[249] prostaglandin E2,[250] 30kD survival factor, taurine, and vitamin A. Other useful cofactors include symptom-alleviating cofactors, including antiseptics, antibiotics, antiviral and antifungal agents and analgesics and anesthetics.

The IPMC therapeutics also may be associated with means for targeting the IPMC therapeutics to a desired tissue. Alternatively, an antibody or other binding protein that interacts specifically with a surface molecule on the desired target tissue cells also may be used. Such targeting molecules further may be covalently associated to the IPMC therapeutic, e.g., by chemical crosslinking, or by using standard genetic engineering means to create, for example, an acid labile bond such as an Asp-Pro linkage. Useful targeting molecules may be designed, for example, using the simple chain binding site technology disclosed, for example, in U.S. Patent No. 5,091,513.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner. For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane,

dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt. Other suitable delivery systems include microspheres which offer the possibility of local noninvasive delivery of drugs over an extended period of time. This technology utilizes microspheres of precapillary size which can be injected via a coronary catheter into any selected part of the body, e.g., the eye, or other organs without causing inflammation or ischemia. The administered therapeutic is slowly released from these microspheres and taken up by surrounding tissue cells (e.g. endothelial cells).

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and

include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For topical administration, the oligomers of the invention are formulated into ointments, salves, gels, or creams as generally known in the art. A wash solution can be used locally to treat an injury or inflammation to accelerate healing.

In clinical settings, a gene delivery system for the therapeutic IPMC gene can be introduced into a patient by any of a number of methods, each of which is familiar in the art. For instance, a pharmaceutical preparation of the gene delivery system can be introduced systemically, e.g., by intravenous injection, and specific transduction of the protein in the target cells occurs predominantly from specificity of transfection provided by the gene delivery vehicle, cell-type or tissue-type expression due to the transcriptional regulatory sequences controlling expression of the receptor gene, or a combination thereof. In other embodiments, initial delivery of the recombinant gene is more limited with introduction into the animal being quite localized. For example, the gene delivery vehicle can be introduced by catheter[251] or by stereotactic injection.[252] An IPMC gene, such as any one of the sequences represented in the group consisting of SEQ ID NOS:1-3, or a sequence homologous thereto can be delivered in a gene therapy construct by electroporation using techniques described.[253]

The pharmaceutical preparation of the gene therapy construct or compound of the invention can consist essentially of the gene delivery system in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle or compound is imbedded. Alternatively, where the complete gene delivery system can be produced intact from recombinant cells, e.g., retroviral vectors, the pharmaceutical preparation can comprise one or more cells which produce the gene delivery system.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

The present invention is further illustrated by the following examples which should not be construed as limiting in any way. The contents of all cited references (including literature

references, issued patents, published patent applications as cited throughout this application) are hereby expressly incorporated by reference.

The practice of the present invention can employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature.[129,174,195,254,255,2256,257,258,259,260,261,262,263,264,265]

EXAMPLES

1. Cloning and Analysis of Human IPM150 and IPM200 cDNAs

a. *Tissues*: Human eyes were obtained from MidAmerica Transplant Services (St. Louis, MO) and the University of Iowa Lion's Eye Bank (Iowa City, IA) and processed within three hours of cardiac cessation under approved Institutional Review Board protocols. Human eyes were obtained from donors less than 60 years of age and without any known ocular disorder. Monkey eyes and other organs were obtained from cynomolgus macaque monkeys (*Macaca fascicularis*) immediately after barbiturate-induced euthanasia. All animals were treated in conformity with the NIH "Guide for the Care and Use of Laboratory Animals", the ARVO "Resolution for the Use of Animals in Research" and the established guidelines of St. Louis University and the University of Iowa.

b. *Isolation of IPM*: Human and monkey neural retinas were separated from the RPE and incubated in 10 mM phosphate buffered saline (PBS), pH 7.4, containing a cocktail of protease inhibitors (2 mM phenylmethylsulfonylfluoride, 10 mM N-ethylmaleimide, 1 μg/ml pepstatin A, 1 μg/ml leupeptin, 0.02% sodium azide, 100 KIU/ml aprotinin, 100 mM 6-amino-n-caproic acid, 5 mM benzamidine-HCl and 0.04% EDTA) for 5-10 min, with slight agitation, to remove soluble IPM constituents (15). The isolated retinas were then placed in (1) 4 M urea in 10 mM PBS containing 0.5% NP-40 and protease inhibitors or (2) distilled water for 7-10 minutes at 4°C, with slight agitation, until the IPMC protein dissociated from the photoreceptor cells. The resulting sheets of insoluble IPMC proteins were isolated, dialyzed against 10 mM PBS pH 7.4, their protein content determined using a micro BCA assay (Pierce), pelleted by centrifugation and frozen at 80°C, or lyophilized for subsequent analyses. Alternatively, pellets were resuspended in PBS and repelleted 7-10 times.

c. *Identification of IPM150 and IPM200 Proteins*: Aqueous-insoluble human and monkey IPMC protein preparations were homogenized and digested with protease-free chondroitin ABC lyase (E.C. 4.2.2.4) (Seikagaku Ltd., Rockville, MD) at a concentration of 2 U/ml, in the presence of protease inhibitors, for 2 hrs. at 37°C. Western blot analyses of chondroitinase-treated IPMC preparations from pig, monkey and human retinas demonstrated bands of approximately 150 kDa and 200 kDa (and minor bands of 180 kDa and 220 kDa in pig and monkey, respectively) that were labeled by AC6S antibody and PNA. The 150kDa bond was designated IPM150, and the 200 kDa band was designated IPM200. The 150 kDa band appears to contain two glycoproteins; one that binds PNA and migrates to a lower molecular weight following exposure to neuraminidase and O-glycanase, and a second neuraminidase-insensitive component that binds PHA-L. Following complete deglycosylation, AC6S binds to a band of approximately 55-58 kDa. These studies indicate that at least two distinct chondroitin 6-sulfate-containing proteoglycans are associated with CMSs.

Proteins were separated on one- or two-dimensional SDS-polyacrylamide gels under denaturing conditions, transferred to sequencing grade PVDF membranes (Immobilon P, Millipore Corp., Bedford, MA) and briefly stained with 0.1% Coomassie blue. IPM150 bands were identified by comparison to adjacent blot strips and incubated with PNA or AC6S antibody. To identify bands of interest, some lanes were blocked in 50 mM Tris buffer, pH 7.4, containing 1 mM $MgCl_2$, 1 mM $CaCl_2$ (TMC) and 0.2% Tween 20, with or without 2% BSA for 45 minutes, and subsequently rinsed in TMC. The blot strips were exposed to HRP-conjugated lectins or antibodies (followed by HRP-conjugated secondary antibodies) diluted in TMC (1/50-1/500), for one hour, and rinsed in TMC and the HRP-conjugated probes visualized using 4-chloro-1-naphthol and hydrogen peroxide. Some blot strips may also be stained for total protein using India ink (1 μl/ml) in TMC with 0.2% Tween 20. Identified bands, especially IPM150 and IPM200 were used for direct amino acid sequencing or eluted in the presence of CNBR.

For direct amino acid sequence, bands containing IPM150 protein and IPM200 protein were excised and the proteoglycans eluted in 25 mM Tris containing 192 mM glycine and

0.1% SDS using a BioRad electro-eluter. The NH_2-terminal amino acid sequence for IPM150 was determined by Edman degradation. Sequence for every region was obtained at least twice.

Two strategies for generating and isolating CNBR peptides from SDS-PAGE-purified IPM150 and IPM200 proteins can be utilized. A blotting procedure allows one to easily remove SDS from PVDF and allows for reliable elution of PVDF-blotted proteins in high yield. As an alternative approach or when tryptic peptides are required, an "in-gel" digestion procedure was used to generate IPM150 and IPM200 peptides.

CNBR Peptides from Blotted Proteins. PVDF blots were washed with HPLC grade water and stained with 0.05% Coomassie blue in 50% methanol for five minutes. Individual protein bands were cut into 1x2 mm pieces and submerged in 500μl of ice-cold 95% acetone for one hour. The supernatant was removed and the PVDF-bound protein washed with another 500 μl of acetone. The membrane was air-dried, submerged in 100 μl of 70% formic acid, and 10 μl of 70 mg/ml CNBR per 10 μg of blotted protein added. Blots were incubated in the dark for 24 hours, at room temperature, the supernatant removed, the membranes dried, and 100 μl of 40% acetonitrile added to the membranes. Following incubation at 37°C, for 3 hours, the supernatant was removed and pooled with the previous supernatant. Finally, the PVDF pieces were extracted again at 50°C with 100 μl of 0.05% TFA in 40% acetonitrile, the supernatant combined with the two prior extracts, and collectively dried under vacuum.

Chromatographically isolated IPM150 and IPM200 were digested by redissolving dried samples in 70% formic acid (v/v) and adding a 20- to 100-fold molar excess (over the methionyl residues) of CNBR, for 24 hours. The reaction was stopped by drying the sample under vacuum.

CNBr Peptides From "In-gel" Tryptic Digestion: Other amino acid sequences were obtained from fragments of IPM150 generated by an "in-gel" trypsin digestion protocol. Briefly, chondroitinase ABC-treated IPM protein preparations were separated by SDS-PAGE, the gels were stained with 0.1% Coomassie and IPM150 was identified and excised. Gel strips were incubated, at 37°C, for 24 hours, in a 1:25 (weight-to-weight) ratio of trypsin (Boehringer Mannheim, Indianapolis, IN) to protein. Following incubation, the gel pieces were shaken at room temperature for 8 hours to extract peptides. The resulting fragments were fractionated using reverse phase HPLC in 0.2 ml of a 2 M urea solution. Separated peptides

were detected by absorbance at 210 nm and collected using a peak detector. Selected peptides were subjected to amino acid sequencing at the W. K. Keck Foundation (New Haven, CT).

Alternatively, IPM150, IPM200, and homolog proteoglycans of the invention can be selectively isolated from other insoluble IPMC glycoproteins based on their high concentration of carboxyl and sulfate ester groups and high negative charge. Insoluble IPMC protein preparations are dialyzed against 8 M urea in 50 mM Tris-HCI, pH 6.8, and subsequently applied to a HPLC TSK DEAE-5PW analytical ion exchange column. Bound proteoglycans are eluted by a linear salt gradient of 0.15-1.0 M sodium chloride. IPM150 and IPM200 elute between 0.4-0.5 M sodium chloride. IPM150- and IPM200-containing fractions are pooled, dialized against distilled water, and lyophilized. Lyophilized samples are chondroitinase ABC-treated and separated by HPLC size exclusion chromatography on a TSK 4000PW-3000PW column in the presence of 4 M guanidine hydrochloride in 10 mM Tris-HCI, pH 7.0, containing protease inhibitors. The two peaks containing IPM150 and IPM200 are collected and the presence of IPM150 and IPM200 in the fractions verified by Western blot analysis. These fractions can be utilized for analysis of core proteins and carbohydrates of IPM150 and IPM200 and homologs, allelic variant and mutants thereof. Other fractions can be saved for analyses of additional IPMC protein constituents.

As an initial step in determining whether the difference in molecular weights of IPM150 and IPM200 is manifested by differences in their protein and/or their oligosaccharide compositions, CNBR and/or tryptic peptides derived from human IPM150 and IPM200 can be separated and the resultant profiles from the two proteoglycans compared for consistent differences in the numbers or sizes of resultant peptides. The resulting peptides are separated on a Vydac C-18 reverse phase HPLC column equilibrated in 0.05% TFA and 1.6% acetonitrile, and eluted with increasing concentrations of the same buffer. Peptides are collected by peak with the use of an ISCO peak detector. Samples are stored at 5°C or applied directly onto polybrene coated GF-C filters and their amino acid sequences determined. This procedure provides"CNBR peptide maps" of IPM150 and IPM200. Differences in numbers of peptides might indicate variation in amino acid sequence, since CNBR cleaves at unoxidized methionine residues. Overall differences in apparent sizes of peptides indicate either a true difference in amino acid composition or the presence of glycanated variants of the same peptide.

The N-termini of CNBR- and/or tryptic-derived peptides can be subjected to amino acid sequencing using an Applied Biosystems sequencer equipped with on-line HPLC systems. Approximately 0.05-1.0 nm of protein/peptide is sufficient to sequence between 15 and 40 residues, respectively. Amino acid sequences obtained in this fashion are verified by sequencing the same peptide at least twice. Since greater than 80% of eukaryotic proteins have blocked amino termini that preclude direct amino acid sequencing, this strategy provides information pertaining to partial amino acid sequences of IPMC core proteins that might be unattainable otherwise. Results obtained from these analyses provide information about differences in amino acid sequences of IPMC core proteins in humans; similar analyses can be performed using monkey and pig IPMC protein if their CNBR profiles differ significantly from those observed in humans.

Direct amino-terminal amino acid sequencing resolved 31 and 20 amino acids of the amino termini of monkey and human IPM150, respectively (Figures 4 and 5, SEQ ID NOS:2 and 4). Alignment of these two sequences reveals a high degree of sequence conservation between the two species (Figure 3, SEQ ID NOS:2 and 4). Comparison of these sequences to those in the NCBI database indicates that IPM150 is not homologous to other proteins.

N-terminal amino acid sequences of eight core proteins from human, monkey and pig IPMC protein were determined. The sequences show that (1) the core proteins of the IPM150, and IPM200 proteins share a strongly conserved N-terminus; (2) that these proteins are conserved among higher mammalian species; and (3) that the amino acid sequences are unique. Specific conservative amino acid substitutions were observed, however, at both the intra- and inter-species levels.

d. *Reverse Transcription-Polymerase Chain Reaction (RT-PCR)*: Total retinal RNA was isolated from monkey and human retinas using RNAStat-60 reagent (Tel-Test, Inc., Friendswood, TX) and 100 ng/tube were reverse-transcribed using random hexamer primers and the GeneAmp® RNA PCR Kit (Perkin Elmer Cetus, Foster City, CA). For the initial experiments, monkey retinal cDNA was amplified for 25 cycles using a degenerate sense PCR primer (5'-TATTAGGAATTCCATYTTYTTYCCIAAYGG-3') (SEQ ID NO:10), designed using the amino acid sequence of the monkey IPM150 amino-terminus and a degenerate antisense primer (5'-TTICCIGCIAGYTCYTGRTARTAIGG-3') (SEQ ID NO:11) which was designed based on the sequence of tryptic peptide #70, derived from human IPM150 (see

Figure 5). During the synthesis of these primers, inosine residues were used in positions of complete degeneracy. In the first step, 200 ng/tube total RNA is reverse transcribed with M-MLV reverse transcriptase using random hexamers or oligo (dT) primers for 5 minutes at room temperature, followed by incubation at 42°C for 15 minutes. The enzyme is then denatured by heating at 95°C for 15 minutes. cDNA is amplified at annealing temperatures 2°C below the melting point of the primers or, in the case of degenerate primers, 2°C below the lowest melting point possible for that primer, for 30 cycles. Ten percent of the reaction is analyzed by agarose gel electrophoresis. If no amplification product is observed, a small aliquot of the reaction is used as template in a second amplification reaction. If an amplification product is present, the experiment is repeated, at least twice, from a new stock of RNA; only those amplification products which occur consistently are used for subcloning. Following amplification, 10% of the reaction is analyzed by agarose gel electrophoresis. When an amplification product is obtained, the experiment is repeated, at least twice, from a new stock of RNA. A monkey amplification product of 580 bp was isolated and ligated into the SrfI site of the vector PCR-Script SK (Stratagene, La Jolla, CA). The resulting clone, designated 70-1, was sequenced and shown to encode 192 amino acids of monkey IPM150, starting at the presumed amino terminus. The deduced amino acid sequence features two possible N-glycosylation sites, numerous potential glycosylation sites, four cysteine residues, and two hyaluronate-binding motifs.

Another clone, designated 9-5#3, was generated by RT-PCR. Human retinal RNA was reverse transcribed, using a primer complementary to nucleotides 2927 to 2943 of the human IPM150 sequence (see SEQ ID NO:12). The resulting single-stranded cDNA was amplified using sense (nucleotides #109-123) (SEQ ID NO:13) and antisense (nucleotides #2675-2689) (SEQ ID NO:14) primers corresponding to sequences flanking the predicted open reading frame of human IPM150. The resulting PCR fragment was ligated into the vector PCRII (InVitrogen, Carlsbad, CA) and subcloned according to standard art known methods.

e. *cDNA Libraries*. Five cDNA libraries, derived from poly $A^{(+)}$ RNA isolated from human and pig ocular tissues, were generated. They are: human retina--$1.6x10^6$ independent clones (random primed; λgt11); pig retina--$0.5x10^6$ independent clones (random- and oligo (dT)-primed; λgt11); and pig retina/RPE/choroid--$5.0x10^6$ independent clones (random- and oligo (dT)-primed; λZapII).

Construction of λZapII phagemid libraries is performed as follows: Double-stranded cDNA with cohesive *Eco*R I ends are synthesized using the SuperScript™ Choice System (Gibco BRL). mRNA is reverse transcripted in a reaction primed by oligo(dT) and/or random hexamers catalyzed by SuperScript™ II out in the same tube to increase the yield of double-stranded cDNA. *Eco*R I adaptors are added without having to methylate or restriction enzyme digest the cDNA. The cDNA is made vector-ready by passage over a column that removes unincorporated adapter-arms (Pharmacia). The mixture is packaged with Gigapack packing extract (Stratagene) and transfected into *E.coli* XL-1 Blue cells (Stratagene). The number of independent clones is assessed prior to the screening of each library.

f. *Screening cDNA Libraries with DNA Probes*. Human and monkey cDNA libraries in λZapII are plated to a density of approximately 3000 pfu per 100 mm plate on *E.coli* XL-1 Blue host cells. Duplicate plaque lifts are prepared using Hybond N nylon filters (Amersham). cDNA probes are ^{32}P-labeled in a random-primed reaction using a multi-primer DNA labeling kit (Amersham), purified over G-50 columns, heat denatured and added to the prehybridization mix. Hybridization can be performed for 8-20 hours at 65°C. After hybridization the filters are washed at room temperature for 3x5 minutes in 2X SSC containing 0.3% SDS and then 2x3 minutes in 0.2X SSC containing 0.3% SDS. Final wash stringencies may be decreased or increased depending on the source and sequence of the cDNA probe and level of background radiation detected on the autoradiograms. Autoradiograms are aligned with the filters and plates. Only plaques giving a signal on both filters are isolated, replated at a lower density (~500pfu/plate) and rescreened until well-isolated positive plaques are obtained. Lambda gtll libraries are screened in a similar fashion, with the exception that *E.coli* *Y*1090 are used in place of XL-1 Blue cells as host.

g. *Isolation of Human IPM150 cDNA Clones*: A commercially available human retinal cDNA library in λgt11 (Clontech Laboratories Inc., Palo Alto, CA) was screened by plaque hybridization, using probes derived from clone 70-1. Plaques were plated to a density of approximately 10,000 pfu per 150 mm Petri dish on *E. coli* Y1090 host cells. Duplicate plaque lifts were prepared using nitrocellulose filters (Schleicher & Schuell, Keene, NH). cDNA probes were labeled with ^{32}PdCTP in a random-primed reaction and hybridized to the plaques for 12-20 hrs at 65°C. After hybridization, the filters were washed at room temperature for 10 min in 2X SSC (0.3 M NaCl and 30 mM sodium citrate) containing 0.3%

SDS, and two additional times, for 10 min each time in 0.2X SSC containing 0.1% SDS at 60°C. Plaques giving signals on autoradiograms derived from both filters were isolated, eluted in λ-buffer (10 mM Tris, pH 7.5 and 10 mM MgCl), replated at a lower density (~500 pfu/plate), and rescreened until isolated, positive plaques were obtained. cDNA inserts of purified λgtll clones were PCR-amplified using primers to the flanking regions of λgtll and ligated into pCRII using the TA cloning kit (InVitrogen, Carlsbad, CA).

Three cDNA clones, designated 8.1.2, 11.1.1 and 12.3.1, were isolated, sequenced and assembled into one contiguous sequence (Figure 5, (SEQ ID NO:3)). The assembled nucleotide sequence of 3,261 bp contains an uninterrupted open reading frame of 2,313 bp and several hundred bp of untranslated 5' and 3' regions. Verification that the assembled sequence encodes human IPM150 is provided by the presence of the complete amino-terminal amino acid sequence, as well as all internal peptides, within the deduced amino acid sequence.

Clone 8.1.2 lacks a 234 bp segment within its 5' region that encodes the amino terminus of IPM150. This clone may represent a cloning artifact or it may indicate the existence of IPM150 isoforms. PCR product encompassing the entire open reading frame of IPM150 was cloned. PCR amplification of reverse-transcribed human retinal RNA yields a 2579 bp fragment which has been subcloned and designated 9p#3. This clone is completely homologous to the assembled sequence and includes the 234 bp region that is absent in clone 8.1.2 (see Figure 5; SEQ ID NO:3).

Clone 12.3.1 encodes a 809 bp fragment of human IPM150, comprising 280 bp of the 5'UTR and 589 nucleotides of the coding region, and shares 94% homology to the 5' region of the monkey homolog between nucleotides 73 and 256 (human) and 2 and 185 (monkey), using BESTFIT analysis with a gap weight of 12 and a length weight of 4. Clone 12.3.1 appears to display the same features/motifs as p70-1 with the exception that an additional N-glycosylation site is present and a potential hyaluronan-binding motif occurs in the leader sequence.

Database searches of the compiled nucleotide and amino acid sequences show that they are novel. A few expressed sequence tags (ESTs), derived from human retina, vein endothelial cells, and brain, however, share homology with IPM150 (GenBank Accession numbers H38604, W26960, H38594, AA326863 and AA296278). Translation of the open reading frame of the assembled cDNA sequences encodes a protein of 771 amino acids with an

isoelectric point of 4.70 and a predicted molecular weight of 86.36 kDa. The protein is generally hydrophilic[70] except for the first 18 amino acids which form a hydrophobic region flanked by charged amino acid residues; this region may represent a signal sequence (Figure 4).

A distinct distribution of consensus sequence sites for N- and 0-linked glycosylation are present in IPM150. There are several consensus sequences for N-linked glycosylation[71] sequestered in the amino-terminal (amino acids 71 to 217 (SEQ ID NO:4)) and the carboxy-terminal (beginning at amino acid 591 (SEQ ID NO:4)) portions of the IPM150 core protein. In contrast, the central domain of IPM150 (between amino acids 220 and 565 (SEQ ID NO:4)) features 17 sites which are suitable for O-linked glycosylation, as predicted by a proposed algorithm for the activity of polypeptide N-acetylgalactosaminyl-transferase.[72] There is almost no overlap between the regions containing potential O- and N-linked glycosylation sites. Hyaluronan binding consensus sequences[73] and several cysteine residues are also present in the amino- and carboxy-terminal regions of the core protein. The distribution of the carboxy-terminal cysteine residues closely resembles that of EGF-like domains,[74] motifs that are present in many extracellular matrix proteins. Consensus sequences for other domains that are commonly associated with extracellular matrix proteins, including immunoglobulin-like[266] and lectin-like domains,[267] are not present in IPM150.

h. *Subcloning cDNA Inserts*. Positive plaques are picked and resuspended in 0.5 ml SM (.1 M NaCl, 1 mM $MgSO_4$, 5 mM Tris, pH 7.5, and 2% gelatin). Inserts in λZapII phagemids are subcloned by *in vivo* excision as described by the manufacturer (Strategine). The resulting pBluescript SK(-) plasmids are transferred into *E.coli* XL-1 Blue. Inserts contained within λgt11 are isolated using the Wizard Lambda Prep kit (Promega), released by *Eco*R1 digestion and subsequently ligated into pBluescript. Plasmid DNA is isolated using Wizard miniprep resins (Promega) and assayed for insert size by restriction enzyme digest and size fractionation on 0.7-1.5% agarose gels.

i. *Anchored PCR*. cDNA amplification using one-sided, or "anchored," PCR can be used to obtain missing sequence information. Regions upstream (5') or downstream (3') of known sequence can be amplified by this method. To amplify unknown sequences downstream, anchored PCR is performed using 100-300 ng total poly $(A)^+$ retinal RNA and these primers: an oligo (dT) primer and two sequence-specific primers, one of the original

amplification reaction (primer 1) and the second (primer 2; an internal sequence-specific primer, which can partially overlap primer 1) for the reamplification reaction. The mRNA is reverse-transcribed as described above, with an oligo (dT) primer. The resulting cDNA is amplified using oligo (dT) and sequence-specific primer 2. Analysis of an aliquot of the first amplification reaction by agarose gel electrophoresis should reveal a smear, whereas analysis of the reamplification reaction should appear as a single band which are subcloned into a plasmid vector, pCRII, using a TA cloning kit version 2.2 (Invitrogen) and sequenced. Amplification of unknown sequences upstream of known sequence by anchored PCR employs a slightly different strategy than for amplification of unknown downstream sequences. To obtain unknown upstream sequences, the reverse transcription reaction is anchored by one of the sequence-specific primers. The resulting cDNA is then modified or "tailed" in a terminal deoxynucleotidyltransferase catalyzed reaction at the 3' end by the addition of a poly $(A)^+$ tail. Two PCR reactions mediated by two sequence-specific primers and oligo (dT) complementary to the newly synthesized tail is carried out as above to yield the desired unique product.

Subcloned cDNA fragments were manually sequenced by dideoxy nucleotide chain termination[268] using the Sequenase Version 2.0 DNA sequencing Kit (Amersham, Arlington Heights, IL) and α-^{35}S-dCTP. Both strands are analyzed at least twice using either vector-specific primers or custom oligonucleotide primers.

j. *Comparisons Between Nucleotides or Amino Acid Sequence:* Comparisons between the obtained sequences and those in the NCBI databases were executed using the BESTFIT software with parameters for gap weight = 7-50 and length weight = 1-4.

2. Characterization of IPM150 and IPM200 Oligosaccharides. In order to characterize potential variations between IPM150 and IPM200 oligosaccharides, to determine the types of oligosaccharides present, and to determine the specific regions of the core glycoproteins that are glycosylated in normal and diseased tissue, CNBR peptides of normal human and monkey IPM150 and IPM200 proteins are generated and their oligosaccharides characterized.

a. *CNBR Maps/Dot Blots*. CNBR peptides are generated from human and monkey IPM150 and IPM200 proteins. Based on the amino acid sequences obtained from the amino-termini of each peptide, the position of the peptide within the core protein are

identified, assuming the complete deduced amino acid sequences are available from cDNAs encoding IPM150 and IPM200. For any given peptide that contains a putative oligosaccharide or GAG attachment consensus sequence, one can determine whether that site is glycosylated.

In order to determine which CNBR peptides are glycosylated, CNBR peptides are incubated with various glycosidases (and/or other agents, such as nitrous acid, that remove oligosaccharides) in the presence of a protease inhibitor cocktail and separated by reverse phase HPLC. The resulting profiles are compared to profiles derived from non-deglycosylated CNBR peptides. A shift in any individual CNBR-derived peptide confirms that it is glycosylated. In addition, this provides information about the type of oligosaccharide and linkage involved. Subsequently, 2-4 µl of CNBR peptides in 0.1% TFA are applied to ProBlott membranes (Applied Biosystems) prewetted with methanol, followed by water, and placed on water-soaked Whitman 3M paper. Membranes are air-dried, rewetted in methanol and subsequently soaked in 15 ml of blocking buffer (50 mM Tris-HCI containing 0.5 M NaCl, and 2% polyvinylpyrrolidone-360), pH 7.5, for one hour at room temperature. The membranes are incubated for 1-2 hours, at room temperature, in blocking buffer containing 10 µl/ml of various HRP-conjugated lectins (including PNA, WGA, LFA, and Con-A) or antibodies (e.g., AC6S) followed by the appropriate HRP-conjugated secondary antibodies. In order to confirm that specific CNBR peptides are glycosylated, they are incubated in the presence of glycosidase(s) (including O-glycosidase, N-glycosidase, chondroitinase, heparanase, neuraminidase), reapplied to ProBlott, reprobed with the same lectins and/or antibodies, and compared to their nondeglycosylated precursors. These blots are utilized to determine the specificities of various monoclonal and polyclonal antibodies for core protein and/or carbohydrate epitopes.

b. *Fluorophore-Assisted Carbohydrate Electrophoresis*. Oligosaccharides associated with IPM150 and IPM200 are profiled, the monosaccharide compositions of specific oligosaccharides (or of the collective oligosaccharide composition of IPM150 and IPM200) are determined, and specific oligosaccharides are isolated, purified, and sequenced. This is accomplished using fluorophore-assisted carbohydrate electrophoresis (FACE), a recently developed technology based on the separation of fluorophore-derivatized carbohydrates on polyacrylamide gels (Glyko; Novato, CA).

c. *Release of Oligosaccharides*. Oligosaccharides are released from intact IPM150 and IPM200 and/or from CNBR peptides. O-linked sugars are released with O-glycosidase, N-linked oligosaccharides with endoglycosidase H or N-glycosidase F, and GAGs with nitrous acid. These methods preserve a free reducing end on the oligosaccharides that are labeled by reductive amination with the fluorescent tag, 1-aminoaphthalene-3, 6, 8-trisulfonic acid (ANTS). Once labeled with ANTS, all carbohydrates acquires a net negative charge that allow their separation on polyacrylamide gels. Pre-packaged kits for all these methods are used as supplied by Glyko (Novato, CA).

d. *Separation of Oligosaccharides*. Oligosaccharides (native or sialyated oligosaccharides following removal of sialic acid) as large as 100 kDa are separated based on size, in comparison with a mixture of dextran standards (or used unsaturated oligosaccharides for GAGs). Since the separation of oligosaccharides is influenced by the charge/mass ratio of the saccharide as well as its hydrodynamic volume, oligosaccharides that contain sialic acid migrate faster than larger neutral oligosaccharides. These separations provide information pertaining to the degree and type of glycosylation on individual CNBR peptides.

e. *Monosaccharide Composition Analyses*. Individual oligosaccharides are eluted from polyacrylamide gels and hydrolyzed into free monosaccharides using acid hydrolysis. Monosaccharides are labeled with ANTS and separated electrophoretically on 5% gels. The resulting profiles of neutral, amine and sialic acids is compared to monosaccharide standards. Gels are photographed using an electronic imaging system based on a CCD camera. Quantitative analyses of the resolved components is obtained by high-resolution, computer-assisted image analysis.

f. *Oligosaccharide Sequencing*. Oligosaccharides eluted from gels are exposed to a battery of specific glycosidic enzymes (including neuraminidase, β-galactosidase, hexosaminidase, α--mannosidase, β-mannosidase, and fucosidase) and rerun on polyacrylamide gels. Migration patterns of variously treated oligosaccharides provide sequence information of specific oligosaccharides. Should more specific data on linkage analysis be required, purified IPM150 and IPM200 and/or CNBR peptides derived from these proteoglycans can be sent to the Complex Carbohydrate Research Center (Athens, GA) for analyses using gas chromatography-mass spectrometry, fast atom bombardment/mass spectrometry or nuclear magnetic resonance.

3. Characterization of Foveal IPMC Protein Glycoconjugates. Western blot analyses was used to characterize the size(s) and carbohydrate composition(s) of fovea-associated IPM glycoproteins and proteoglycans. 1-1.5 mm diameter punches of monkey and human foveas were homogenized in PBS containing protease inhibitors. A portion of the homogenate was digested with chondroitinase ABC and another portion completely deglycosylated. The enzyme-treated and untreated portions were separated by SDS-PAGE and electroblotted onto nitrocellulose. The blots were probed with antibodies generated against components of the extrafoveal IPM, including IPM150 and IPM200. Reaction of these antibodies with fovea-derived bands of the same apparent molecular weight(s) as IPM150 and IPM200 indicate that the fovea contain the same, or similar, proteoglycans present in the extrafoveal IPMC protein. In addition, the identification of IPM150 and IPM200 core glycoproteins in the lanes from non-chondroitinase-treated foveas provided evidence that foveal IPMC core proteins are similar to IPM150 and/or IPM200, except for the presence of chondroitin sulfate. Blots were also incubated with various lectins that bind to foveal IPMC protein. Various bands were excised and their N-terminal amino acid sequences determined to assess sequence homology with IPM150 and IPM200. Alternatively, previously undescribed glycoproteins were identified on these blots using probes that bind foveal IPMC protein. In this case, amino acid sequences of the proteins in these bands were obtained to determine whether they are unique or whether they are glycanated variants of IPM150 and IPM200. Should they possess the same core protein, but different carbohydrates, the CNBR strategy can be employed to characterize their oligosaccharides. Concurrent with the biochemical studies, sections of normal human foveas were hybridized with various probes, including IPM150 and IPM200 DNA probes, to determine whether messages encoding IPM150 and/or IPM200 are expressed in foveal cones.

4. Characterization of Other Identified Soluble and Insoluble IPMC Protein Constituents. An additional number of insoluble IPMC protein-associated glycoproteins and proteoglycans have been identified in human and monkey retinas. These include distinct 105 kDa and 80 kDa glycoproteins bound by a number of IPM-specific antibodies, a 140 kDa proteoglycan that migrates on gels without prior chondroitinase treatment, and a 150 kDa neuraminidase-resistant PHA-L-binding glycoprotein. These molecules can be characterized

using procedures similar to agarose gel electrophoresis, in combination with Western blotting. Proteins are separated on one- or two-dimensional gels, transferred electrophoretically to nitrocellulose membranes and proteins or CNBR/tryptic peptides derived from them sequenced. Unique proteins are characterized further. The sequence information are then used to design degenerate oligonucleotide probes for RT-PCR of retinal RNA using a strategy similar to hat employed in isolating IPM150-associated cDNA clones.

The HPLC-purified peak that contains 800-900 kDa molecules following chondroitinase treatment can be digested with CNBR to determine whether this high molecular weight IPM components contains glycoconjugates other than IPM150 and IPM200. This is accomplished by comparing the CNBR peptide profiles to those of IPM150 and IPM200. If these profiles suggest the existence of additional molecules, these fragments are collected and subjected to amino acid sequencing to determine potential homology to other molecules. It is also anticipated that this high molecular weight peak may contain hyaluronan; this is based on our hypothesis that IPMC hyaluronan may stabilize the IPMC protein through interactions with CD44 and/or IPM150 and IPM200. In order to determine the presence of hyaluronan in this peak, the peak are analyzed using FACE.

5. Anti-IPMC Antibodies

a. *Generation of Antibodies Directed Against IPM150- and IPM200-Derived Peptides*. Polyclonal antibodies are generated against amino- and carboxy-termini peptides of IPM150 and IPM200. If the core protein amino acid sequences are strikingly similar, it is unlikely that synthetic peptide antibodies distinguish between the two proteoglycans. If this is the case, antibodies are generated against CNBR-derived peptides from IPM150 and IPM200 that share the same, or similar, amino acid sequences, but exhibit significant differences in their oligosaccharide compositions.

b. *Synthetic and CNBr Peptides as Immunogens*. Two general approaches are available for generation of synthetic peptides; these include preparation of antigens from bacterial over expression vectors or synthesis of peptides by solid-phase technology. Both approaches have their inherent advantages and disadvantages. The solid-phase approach has major advantages if the antigen is known to be highly conserved because of the way in which the peptide is displayed to the immune system, since particular regions of protein can be

targeted specifically for antibody production, and because they can be prepared immediately after determining the amino acid sequence. However, the disadvantages are that the resulting antibodies may not recognize the native antigen and that they are more expensive to produce than bacterial fusion protein antigens.

c. *Immunizations and Screening of Antisera*. New Zealand white rabbits and chickens have been utilized for the development of heterologous antisera against IPM150- and IPM200-derived peptides. Synthetic and CNBr-derived IPMC peptides are coupled to keyhole limpit hemocyanin (KLH) using glutaraldehyde as a crosslinker prior to immunization. Some rabbits may be immunized with nonsense peptides of the same size; the resulting antisera are used as controls. Other carriers (including BSA, ovalbumin or PPD) and bifunctional crosslinkers, can be utilized in the event that KLH-conjugated peptides precipitate or do not elicit an immune reaction. As an alternative strategy, multiple antigen peptides (MAPS) can be synthesized on an immunologically inert lysine dendritic core using F-moc chemistry. Anti-peptide polyclonal antibodies with high titer values can be generated using this relatively new technique.

Immunogens are injected subcutaneously (500 µl/site; 10 sites/animal) and/or subscapularly (100 µl/site; 40 sites/animal). Test bleeds from the marginal ear vein (5-10 ml) are made 7-14 days following the second immunization and assayed for specific activity by immunofluorescence and immunoblotting. Rabbits producing relatively high titers of specific antibody activity are continued on "booster" immunizations every 6 weeks. These animals are bled every two weeks and the antisera stored in aliquots at -80°C. Polyclonal antibodies raised against peptides are purified from rabbit anti-serum on a column of the immunogen coupled to thiol-Sepharose 4B (Pharmacia, LKB, Biotechnology). Antibodies are eluted with 0.1 M glycine HCl (pH 2.8) into Tris-buffered saline (0.02 M Tris-HCl, pH 7.4, and 0.15 M NaCl) and dialyzed. Antibody capture assays using purified peptide of origin coupled to a solid phase (e.g. ELISA) can be used to determine antigenicity, to quantitate antibody titers, and to compare epitopes recognized by different antibodies. This approach allows one to determine rapidly whether antibodies distinguish between IPM150- and IPM200- derived or synthetic peptides. Chickens have been immunized using a similar protocol except that IgY is isolated from egg yolks.

Assuming that such antibodies are generated, they can be screened on Western blots and tissue sections to determine their specificities for native IPM150 and IPM200. Controls include the use of preimmune sera and nonsense peptide antibodies. For immunogens coupled to KLH, positive sera are screened in assays using a second glutaraldehyde-coupled peptide since, in some cases, the glutaraldehyde bridge forms a portion of an epitope recognized by the antibody. Antibodies to carrier proteins or to the coupling reagent are removed by affinity purification of the anti-peptide antibodies on columns prepared with conjugates of the peptide to a second carrier molecule. Based on previous experiences, it is anticipated that antibodies of high titer can be produced that distinguish IPM150 from IPM200.

d. *Generation of Monoclonal Antibodies*. Monoclonal antibodies have been generated using an *in vivo* immunization technique known generally in the art. Balb/C mice were immunized with IPM150 and IPM200 peptides. Hybridomas showing high specific antibody concentrations and activity were cloned by means of limiting dilution and the monoclonal antibodies produced employed in immunohistochemical and morphological assays. Some hybridomas were propagated as ascites tumors by injection of these cells into pristane-primed Balb/C mice.

e. *Lectin and Antibody Immunohistochemistry*. Eyecups from normal, immature, aged and diseased humans were fixed in 4% formaldehyde in 100mM sodium cacodylate buffer, for 2-4 hr, and then rinsed for 6 hr, embedded in acrylamide, frozen in liquid nitrogen, and sectioned to a thickness of 5-6 μM. FITC-conjugated lectins or antibodies, followed by the appropriate FITC-conjugated secondary antibodies, are applied to sections as published previously. Competitive haptens, preimmune sera and other appropriate controls are utilized in all studies.

6. Substructural Localization of IPM150 and IPM200 in Cone Matrix Sheaths.

a. The CMS is not only a biochemically distinct component of the IPM, but is structurally distinct as well. Longitudinally-orientated filaments extend the entire length of the CMS and terminate in filamentous rings at its apical and basal ends. The relationship of IPM150 and IPM200 to the longitudinal filaments of CMSs and elucidation of their topographical association with VnR at the surface of cone inner segment ellipsoids and rod photoreceptors can be determined as follows.

Retinas were dissected from human and monkey eyes and rinsed in ice cold 10mM PBS, pH 7.4, containing 1 mM $CaCl_2$ and 1 mM $MgCl_2$ and placed in ice-cold distilled water containing 2mM $CaCl_2$ to remove sheets of insoluble IPM. The isolated IPMC preparations were then transferred to calcium-free distilled water, which resulted in expansion of CMSs, thereby facilitating visualization of their substructural organization. IPMC protein preparations were fixed in 4% paraformaldehyde, in some cases. These preparations were subsequently exposed to one or more FITC-conjugated lectins, IPM150/IPM200 synthetic peptide-and CNBr-derived antibodies, and/or other IPM-binding antibodies, followed by exposure to the appropriate secondary antibody conjugated to FITC or rhodamine. Appropriate controls were utilized to confirm specificity. The IPMC protein preparations were visualized by epifluorescence or confocal microscopy to determine the distribution of antibody-recognized epitopes. Confocal microscopy using a BioRad MRC-600 laser scanning confocal imaging system permitted visualization of thin optical sections; a Z-series of serial confocal images was recorded and computer assembled to give a three-dimensional image of the labeled IPM. Kalman averaging of confocal images was performed with at least 20 images taken at each level and images analyzed using COMOS, a windows-driven software program. If higher resolution is required, electron microscopic immunocytochemistry can be performed.

b. *In Situ Hybridization*. *In situ* hybridization was used to determine the cellular sources of specific mRNAs encoding IPMC protein molecules. Using the Riboprobe®System (Promega), ^{35}S-labeled anti-sense and sense RNA probes were synthesized from a 605 bp fragment of the human IPM150 cDNA, spanning from bases 1636 to 2241 in the presence of ribonuclease inhibitor and the appropriate primer (T3 or T7 depending on the desired template). Eyes and other tissues were fixed in 4% paraformaldehyde (+/- 0.5% glutaraldehyde), dehydrated in ethanol and embedded in diethylene glycol distearate. 1 μm thick tissue sections were cut, the embedding medium removed with toluene, and the tissue rehydrated with decreasing amounts of ethanol in DEPC-treated water. Sections were then incubated in 0.1 M triethanolamine and 0.25% acetic anhydride, pH 8.0, subsequently rinsed in 2X SSC and water, and then air-dried. Hybridization was conducted in 15 μl hybridization buffer composed of 50% formamide containing 10mM Tris-HCl, 0.3M NaCl, pH 8.0, 1 mM EDTA 10% dextran sulfate, 5 μg/ml yeast tRNA, 100 mM DTT, 1X Denhardt's solution and

$4x10^5$ dpm of probe, at 42°C, for 14-16 hr. Evaporation was prevented by covering the sections with a NaOH-washed coverslip and sealing the edges with rubber cement. A negative control, consisting of the sense probe, was present on an adjacent tissue section on each slide. After hybridization, the coverslips were removed and unhybridized probe rinsed off using several washes of 4X SSC, followed by a treatment with DNase free RNase A (10μg/ml). The slides were washed twice for 15 min in 2X SSC, at room temperature, and three times for 15 min in 0.1X SSC at 45°C followed by dehydration in increasing concentrations of ethanol. The slides were dipped in a 1:1 dilution of Kodak NTB-2 emulsion and stored in a light-tight box at 4°C. At various exposure times, slides were developed in D19, stained with 1% toluidine blue and visualized by light microscopy. Provided section thickness is constant and that control and experimental slides are processed at the same time and in replicates, quantitative date (grains/μm^2 of tissue) can be generated, if necessary.

c. *In Situ RT-PCR*. To distinguish between various populations of cells (e.g. rods versus cones) in developing, aging and/or pathologic retinas, *in situ* RT-PCR can be utilized.[269,270] Fresh tissue can be fixed in 4% paraformaldehyde in PBS, washed in PBS, dehydrated in a series of graded ethanols and embedded in paraffin.[271] 3μm sections are cut and mounted on silanized slides, paraffin is removed at 55°C and the tissue is rehydrated. The sections are digested in a solution of 1 μg/ml Pronase in 50 mM Tris-HC1, pH 7.5, and 5 mM EDTA for 30 min at 37°C and then washed. DNA is degraded by an overnight incubation in RNase-free DNase in the presence of RNase inhibitor and the solution removed by washing the slides in a large volume of 50 mM Tris-HC1, pH 7.5.

RNA exposed on sections is reverse transcribed in 10 μl of a buffer containing 5 mM KC1, 1 mM Tris-HC1, pH 8.3, 5 mM MgCl, 10 mM dNTPs, 1 ng downstream primer, 1μl RNase inhibitor, 8 μl DEPC-treated water and 20U M-MLV reverse transcriptase. The specimens are covered with Parafilm and incubated in a moist chamber, at 42°C, for one hr. The RT mixture is then replaced with a PCR mixture including 100 ng PCR primers and 1 μl α^{35}S-dATP. The specimens are then coverslipped and heated to 75°C. Once this temperature is achieved the coverslip is removed and a preheated aliquot containing 1 μl Taq polymerase added for 20 cycles of amplification. Slides are washed and processed for autoradiography.

To determine absolute differences in the abundance of particular mRNAs in developing, aging or diseased tissues, e.g., retinas, quantitative competitive RT-PCR may be

employed.[270,272] This technique is accurate enough to detect differences in mRNA levels as small as 2-3 fold.

d. *RNA Isolation*. Total RNA was isolated from retinal, other ocular, and extraocular tissues (all major organs) based on the method of Chirgwin *et al.*,[273] except that cesium trifluoroacetate (Pharmacia) was used instead of CsC1, or by using TRIzol (Gibco BRL) in a protocol based on the method of Chomczynski and Sacchi.[274] The resulting pellet was stored at -80°C until required for cDNA synthesis, Northern blot analyses or RT-PCR. When mRNA was required, it was isolated from total RNA using an oligo-dT cellulose-based protocol (Micro Fast-Track-mRNA Isolation kit; Invitrogen Corp.) in which poly $(A)^+$ RNA was bound to oligo(dT) cellulose, washed and subsequently eluted. This method permitted the isolation of 1-5 μg poly (A)+ RNA per human retina. The quality/integrity of RNA obtained was assessed on agarose gels and on Northern blots.

e. *Northern Blot Analyses*: Total RNA and/or poly $(A)^+$ RNA from various ocular and non-ocular tissues was isolated as described above and fractionated on denaturing (formaldehyde-containing) agarose gels. Prior to electrophoresis the samples were denatured by heating at 55°C, for 15 min., in 6.5% formaldehyde and 50% formamide in MOPS running buffer (40 mM morpholinopropanesulfonic acid, pH 7.0, containing 100 mM sodium acetate and 10 mM EDTA). After electrophoresis, the samples were transferred to nitrocellulose or nylon-based membranes by capillary transfer using 20X SSC. The RNA immobilized on the membranes was hybridized with a specific ^{32}P-labeled cDNA probe (clone 70-1-1, corresponding to bp 350-899 of the human cDNA) for detection of the corresponding transcripts and a control actin cDNA to ascertain successful transfer and integrity of the sample.

f. *Tissue Expression*. A commercially available RNA dot-blot containing 100-500 ng poly A^+ RNA from various human tissue sources (Clontech Laboratories Inc., Palo Alto, CA) was probed with a ^{32}P-labeled cDNA corresponding to nucleotides 817-3160 of human IPM150. Hybridization and rinse conditions were exactly as described above.

In order to determine the cellular source(s) of IPM150, Northern and *in situ* hybridization analyses have been conducted. IPM150 cDNA probes hybridized to a 3.9 kb transcript that is present in relatively high abundance in retinal RNA. Occasionally, a larger transcript of approximately 6.5 kb was also detected, albeit at a much reduced signal strength.

No signal was detected on Northern blots of RNA isolated from RPE/choroid, iris or cornea, suggesting that, in the human eye, IPM150 was transcribed only in the neural retina. Dot blot analyses of polyA$^+$ RNA from 50 different adult and fetal human tissues indicate that IPM150 mRNA, or transcripts with a similar nucleotide sequence are present in adult lung, liver, kidney, thymus and small intestine. Weak hybridization to fetal lung and thymus RNA as well as a number of additional adult tissues, was also observed. Distinct hybridization of IPM150 antisense riboprobes to the human retinal outer nuclear layer (ONL) was observed on sections of human retina, RPE and choroid. IPM150 transcripts were present within both rod and cone photoreceptor cells. No labeling of any other region was observed.

7. *In Vivo* Assessment of RGD-Dependent Adhesion System.

a. In order to determine whether adhesion is maintained by an RGD-dependent mechanism involving IPM150, IPM200, and/or other IPMC constituents in normal versus diseased tissue, a variety of blocking antibodies and RGD-containing peptides are injected into the subretinal space of monkeys and pigs. RGD-containing peptides and/or antibodies directed against integrins have been used to inhibit cell adhesion in culture.[275, 276,277,278] Similarly, antibodies to various cell adhesion molecules have been utilized to perturb adhesion in the retina in vivo.[47,51]

Pigs can be utilized for these studies in order to develop the assay and to provide initial information. Provided these initial studies perturb retinal adhesion, the results are confirmed in monkeys. Groups of 4 pigs each are injected subretinally with either 1) RGD-containing peptides known to inhibit VnR-based adhesion (GRGDSP, GRGDTP, GdRGDSP, n-Me-GRGDSP, GRGDSPASSK, and GPenGRGDSPCA) or, 2) blocking VnR antibodies [mouse anti-human αVβ5 (PVF6); mouse anti-human αV (VNR147 and VNR139); rabbit anti-human αVβ3/5]. Initially, F(ab) or F(ab)$_2$ fragments of these antibodies can be prepared and utilized. Peptides (50 μg/ml) or antibodies (1:100) are dissolved in Hanks solution, pH 7.3, and loaded into a micropipette with tip diameter of approximately 50 μm. The micropipette is inserted through a limbal slit and passed across the vitreous until the tip penetrates the central retina. Approximately 5 μl is injected into the subretinal space, creating a small 3-4 mm diameter retinal detachment. A similar bleb is made in another quadrant using Hanks solution

containing non-sense peptides or antibodies or Hanks solution alone. IPMC protein diffusion is confirmed by injecting a ^{14}C-labeled peptide followed by tissue autoradiography.

ERG recordings are made on each of the 4 animals in each group immediately prior to euthanasia at 6 hr, 12 hr, 24 hr and 48 hr following injection and retinal adhesion estimated using an established peeling assay.[158] In brief, eyes are enucleated rapidly (within 15-30 seconds) and the retinas are peeled manually from the RPE in the quadrant containing the retinal bleb, within 30 sec after enucleation. The peeled retinas are fixed for immunohistochemical studies to determine the effects of these peptides and antibodies on IPM structure (especially cone matrix sheaths) and to determine the percentage of retina covered with pigment (to be scored 0-100%, in increments of matrix sheaths) and to determine the percentage of retina covered with pigment (to be scored 0-100%, in increments of 10%, where 100% indicates firmest adhesion). A zone of decreased or no pigment adherence around the original injection sites indicates a loss of retinal adhesiveness beyond the site of injection, providing evidence that the VnR participates in adhesion via an RGD-dependent system. This assay is based on previous observations suggesting that the attachment of cone matrix sheaths to cells of both the RPE and the neural retina is sufficiently strong to result in layers of RPE being separated from Bruch's membrane following manual separation of the neural retina from the RPE within two minutes of death in primates. The basis for this post-mortem loss in adhesiveness is most likely related to physiological dysfunction of an adhesion receptor-ligand system, rather than to degradation of IPM (especially cone matrix sheaths) constituents since no changes in either the structure or distribution of the IPM-associated molecules is observed until approximately 3-4 hours post-mortem in human eyes. The procedures outlined above must be conducted within the first couple of minutes of death to provide meaningful data, since the molecular mechanism that mediates adhesion between the IPM and RPE is extremely sensitive to post-mortem changes.[35] The most straight forward result is one in which the retina easily separates from (or falls off) the RPE in a large zone surrounding the initial injection sight. In addition to the peeling assays, the effects of perturbants on retinal adhesion can be assessed anatomically.[93]

b. *Identification of RGD Consensus Sequences in IPM150 and IPM200*. Once obtained, the complete amino acid sequences of IPM150, IPM200, and other IPMC glycoprotein core proteins can be examined for the presence or absence of RGD consensus

sequences. If an RGD consensus sequence(s) is identified, one can confirm that these proteoglycan core proteins are capable of binding to the VnR associated with photoreceptor and RPE cells. The techniques outlined below can be employed toward this goal.

c. *IPM Peptide Antibodies*. In the event that IPM constituents contain RGD consensus sequences, antibodies can be generated against synthetic peptides containing the RGD sequence and adjacent flanking regions (or, if necessary, CNBr-derived peptides containing the RGD sequence). The specificity of these antibodies for an epitope containing RGD can be assessed using solid phase antibody capture assays. In addition, the binding of these peptide antibodies to native IPMC proteins can be assayed immunohistochemically. Provided these antibodies bind native IPMC proteins, they can be utilized in perturbation assays aimed at assessing the involvement of these constituents in retinal adhesion as outlined in above.

d. *Ligand-Receptor Blotting*. Isolated outer segments and cultured RPE cells (which express VnR) can be extracted for 1 hr, at 4°C, in PBS containing 2mM PMSF, and 200 mM octylglucoside. This lysate is separated by non-denaturing PAGE and transferred electrophoretically to nitrocellulose membranes.[50] These membranes are incubated with biotinylated IPM150 or IPM200 core protein. The blots are washed and incubated with HRP-conjugated streptavidin and color-developed in DAB containing 0.01% hydrogen peroxide. Adjacent lanes are probed with polyclonal antibodies directed against αV and β3/5 subunits and/or αVβ3/5. Binding of IPM150 or IPM200 core proteins to these subunits indicates that IPMC proteins are capable of binding to VnR.

e. *IPM Core Protein Affinity Chromatography*. IPMC protein core affinity columns are prepared by coupling 0.5 mg of proteoglycan core (either deglycosylated native proteoglycan or bacterially-expressed core) to 1 ml Sepharose CL-4B beads after CNBr activation (0.05 g/ml beads). RPE and outer segment lysates are centrifuged at 1000 g for 10 min, their supernatants subjected to the affinity column pre-equilibrated with 50 mM octylglucoside and 1 mM PMSF in PBS, pH 7.4, and eluted with a linear gradient of 0.15-1 M sodium chloride and/or RGD-containing peptides in the presence of $MgCl_2$, $CaCL_2$ or EDTA followed by 0.1 M glycine and 8M urea. A control column is made using BSA coupled to Sepharose CL-4B. Fractions eluted from the column are analyzed on Western blots using antibodies directed against the αV and β3/5 subunits of the VnR. If necessary, RPE cells are

labeled with ^{125}I using the lactoperoxidase-catalyzed iodination for 40 min on ice to label VnR prior to chromatography.

f. *Identification of Other Retinal Adhesion-Related Ligands in the IPMC Protein*. Based on the observation that VnR and CD44 are associated with the plasma membranes of cells bordering the interphotoreceptor space, these adhesive receptors may mediate the linkage of photoreceptor, Müller and RPE cells via are yet unidentified IPMC ligand(s). Since it is known that CD44 interacts specifically with hyaluronan and that the N-terminus of IPM150 contains two potential hyaluronan binding motifs, we speculate that hyaluronan may stabilize the IPMC protein through interactions with CD44, IPM150 and/or IPM200. Such a system provides a framework for a molecular model of retinal adhesion, whereby IPMC proteins are bound to the extracellular domains of VnR and hyaluronan, on the one hand, and to CD44, on the other.

Hyaluronan (hyaluronic acid, hyaluronate), a high molecular mass polysaccharide secreted by many cell types, often plays a dynamic role in cell-extracellular matrix interactions.[84] Immunohistochemical studies of normal, developing and aging primate retinas, employing antibodies directed against hyaluronan (e.g. clone N-DOG1), can conduced to confirm the presence of hyaluronan within the IPM and to identify its distribution. The use of acrylamide embedding is advantageous in these investigations since it avoids the artifactual redistribution of normally hydrated IPMC proteins that occurs following tissue dehydration. The topographical relationship of hyaluronan to other IPMC proteins, including IPM150 and IPM200, can be assessed using isolated preparations of insoluble IPMC protein and the presence of hyaluronan in isolated IPMC protein can be determined biochemically using FACE. In addition, a battery of antibodies directed against various hyaluronan-binding proteins/ receptors and epitopes and other ligands bound by CD44 and VnR can be screened on sections of retina, as they become available.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents of the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

REFERENCES

1. Elman and Fine (1989), The Retina, Ryan, S.J., ed., *CV Mosby*, St. Louis, pp. 175-200
2. Sarks et al. (1980), *Trans. Ophthalmol. Soc. U.K.* 100:414-422
3. Sarks et al. (1994), *Eye* 8:269-283
4. Nathans et al. (1989), *Science* 245:831-838
5. Wells et al. (1993), *Nature Genet.* 3:213-218
6. Nichols et al. (1993), *Nature Genet.* 3:202-207
7. Weber et al. (1994), *Nature Genet.* 8:352-355
8. Capon et al. (1989), *Ophthalmol.* 96:1769-1777
9. Hamilton et al. (1989), *Ophthalmol.* 97:1399
10. Apte et al. (1994), *Genomics* 19:86-90
11. Pennisi, E. (1998), *Science* 281:814-817
12. Gartner and Henkind, (1981), *J. Ophthalmol.* 65:23-28
13. Marshall and Laties (1985), Retinal Degeneration: Experimental and Clinicial Studies, pp. 389-400
14. Yuodelis and Hendrikson (1986), *Prod. Natl. Acad. Sci. USA* 91:5456-5460
15. Dorey et al. (1989), *Invest. Ophthalmol. Vis. Sci.* 30:1691-1699
16. Gao and Hollyfield (1992), *Invest. Ophthalmol. Vis. Sci.* 33:1-22
17. Curcio et al. (1993), *Invest. Ophthalmol. Vis. Sci.* 34:3278-3296
18. Young (1987), *Surv. Ophthalmol.* 31:291-306
19. Young (1992), *JAMA* 84:353-358
20. Organisciak and Winkler (1994), Progress in Retinal and Eye Research, pp. 1-30
21. Sandberg et al. (1994), *Invest. Ophthalmol. Vis. Sci.* 35:2734-2740
22. Grunwald et al. (1993), *Invest. Ophthalmol. Vis. Sci.* 34(13):3609-3613

23. Ramrattan et al. (1994), *Invest. Ophthalmol. Vis. Sci.* 35(6):2857-2864

24. Feeney-Burns et al. (1980), *Am. J. Ophthalmol.* 90:783-791

25. Newsome et al. (1994), Progress in Retinal and Eye Research, pp. 101-124

26. Sarks (1976), *Brit. J. Ophthalmol.* 60:324-341

27. Moore et al. (1994), *Invest. Ophthalmol. Vis. Sci.* (Suppl.)35(4):1502

28. Hogan (1972), *Trans. Am. Acad. Opth. Otolaryngol.* 76:64-80

29. Boulton and Marshall (1986), *Br. J. Ophthalm.* 70:808-815

30. Weiter et al. (1986), *Ophthalmol. Vis. Sci.* 27:145-152

31. Eldred (1987), *Adv. Bios.* 64:23-36

32. Dorey et al. (1993), *Invest Opthalmol. Vis. Sci Suppl.,* 34:1164

33. Taylor et al. (1993), *J. Toxicol. Indust. Health* 9:349-371

34. Sheraidah et al. (1993), *Ophthalmol.* 100:47-51

35. Burns and Feeney-Burns (1980), *Trans. Am. Ophth. Soc.* 78:206-225

36. Bressler et al. (1990), *Arch. Ophthamol.* 108:1442-1447

37. Sarks (1993), Clinicopathological Observations of Drusen. in *Association for Research in Vision and Ophthalmology Annual Meeting*, Sarasota, FL.

38. Killingsworth et al. (1990), *Eye* 4:613-621

39. Gass (1973), *Arch. Ophthalmol.* 90:206-217

40. Piguet et al. (1993), *Br. J. Ophthalmol.* 77:400-403

41. Aszódi et al. (1998), *J. Mol. Med.* 76:238-252

42. Bruckner-Tuderman et al. (1998), *J. Mol. Med.* 76:226-237

43. Hageman and Kuehn (1998), The Pigmented Retinal Epithelium: Current Aspects of Function and Disease, pp. 417-454

44. Hageman and Johnson (1991), *Progr. in Ret. Res.* 10:207-249

45. Hewitt and Adler (1989), *Retina* 1:57-64

46. Mieziewska (1996), *Microsc. Res. Tech.* 15;35:463-471

47. Hageman et al. (1995), *Arch. Ophthalmol.* 113:655-660

48. Hollyfield et al. (1989), *Retina* 9:59-68

49. Johnson and Hageman (1991), *Invest. Ophthalmol. Vis. Sci.* 32:1951-1957

50. Lazarus and Hageman (1992), *Invest. Ophthalmol. Vis. Sci.* 33:364-376

51. Marmor et al.(Date?) *Retina* 14:181-186

52. Yao et al. (1990), *Invest. Ophthalmol. Vis. Sci.* 31:2051-2058

53. Yao et al. (1992), *Invest. Ophthalmol. Vis. Sci.* 33:498-503

54. Yao et al. (1994), *Invest. Ophthalmol. Vis. Sci.* 35:744-748

55. Chu et al. (1992), *Graefes Arch. Clin. Exp. Ophthalmol.* 230:476-482

56. Johnson et al. (1989), *Inherited and Environmentally Induced Retinal Degenerations*, pp. 217-232

57. Lazarus et al. (1993), *Exp. Eye Res.* 54:531-541

58. Porello and LaVail (1986), *Curr. Eye Res.* 5:981-993

59. Seddon and Hennekens (1994), *Arch Ophthalmol.* 112:176-179

60. Tso (1989), *Invest. Ophthalmol. Vis. Sci.* 30:2430-2454

61. Seddon et al. (1994), *J. Am. Med. Assoc.* 272:1413-1420

62. Goldberg et al. (1988), *Am. J. Epidemiol.* 128:700-710

63. Staurenghi (1992), *Invest. Ophthalmol. Vis. Sci.* 33(Suppl):2690

64. Figueroa et al. (1994), *Retina* 14:391-396

65. Litchfield, TM. (1997), *Exp. Eye Res.* 64:655-666

66. Kelsell et al. (1998), *Am. J. Hum. Genet.* 53:274-279

67. Felbor et al. (1998), *Cytogenet. Cell Genet.* 81:12-17

68. Gehrig et al. (1998), *J. Med. Genet.* 35:641-645

69. Wang, et al. (1996), *Mol. Brain Res.* 41:269-278

70. Kyte and Doolittle (1982), *J. Mol. Biol.* 157:105-132

71. Kornfeld and Kornfeld (1985), *Ann. Rev. of Biochem.* 54:631-664

72. Hansen et al. (1995), *Biochem. J.* 308:801-813

73. Yang et al. (1994), *EMBO J.* 13:286-296

74. Rees et al. (1988), *EMBO J.* 7:2053-2061

75. Blanks et al. (1993), *Exp. Eye Res.* 57:265-273

76. Toyama-Sorimachi et al. (1995), *J. Biol. Chem.* 270:7437-7444

77. Bach and Berman (1971), *Biochim. Biophys. Acta.* 252:453-461

78. Korte et al. (1994), *Microscopy Res. Tech.* 29:344-349

79. Tate et al. (1993), *Arch. Ophthtalmol.* 111:963-967

80. Bourdon et al. (1987), *Proc. Natl. Acad. Sci. USA* 84:3194-3198

81. Li et al. (1992), *Arch. Biochem. Biophys.* 296:190-197

82. Huber et al. (1988), *J. Biol. Chem.* 263:752-756

83. Aumailley et al. (1998), *J. Mol. Med.* 76:253-265

84. Ellison et al. (1991), *J. Biol. Chem.* 266:21150-21157

85. Hochuli et al. (1987), *J. Chromatogr.* 411:177

86. Janknecht et al. (1991), *Proc. Natl. Acad. Sci. U.S.A.* 88:8972

87. Current Protocols in Molecular Biology, Ausubel et al., eds., *John Wiley & Sons*, N.Y., 1992

88. Hyrup B. et al. (1996), *Bioorgan. Medicinal Chem.* 4:5-23

89. Perry-O'Keefe et al. (1996), *Proc. Natl. Acad. Sci. USA* 93:14670-14675

90. Finn et al. (1996), *Nucl. Acids Res.* 24: 3357-3363

91. Mag et al. (1989), *Nucl. Acids Res.* 17: 5973-5988

92. Peterser et al. (1975), *Bioorganic Med. Chem. Lett.* 5:1119-11124

93. Van der Krol et al. (1988), *BioTechniques* 6:958-976

94. Stein et al. (1988), *Cancer Res.* 48:2659-2668

95. Wagner (1994), *Nature* 372:333

96. Letsinger et al. (1989), *Proc. Natl. Acad. Sci. USA* 86:6553-6556

97. Lemaitre et al. (1987), *Proc. Natl. Acad. Sci. USA* 84:648-652

98. PCT Publication No. W088/09810, published December 15, 1988

99. PCT Publication No. W089/10134, published April 25, 1988

100. Krol et al. (1988), *BioTechniques* 6:958-976

101. Zon (1988), *Pharm. Res.* 5:539-549

102. Eglom et al. (1993), *Nature* 365:566

103. Gautier et al. (1987), *Nucl. Acids Res.* 15:6625-6641

104. Inoue et al. (1987), *Nucl. Acids Res.* 15:6131-6148

105. Inoue et al. (1987), *FEBS Lett.* 215:327-330

106. See (1988), *Nucl. Acids Res.* 16:3209

107. Sarin et al. (1988), *Proc. Natl. Acad. Sci. USA* 85:7448-7451

108. Bernoist and Chambon (1981), *Nature* 290:304-310

109. Yamamoto et al. (1980), *Cell* 22:787-797

110. Wagner et al. (1981), *Proc. Natl. Acad. Sci. USA* 78:1441-1445

111. Brinster et al. (1982), *Nature* 296:39-42

112. See, e.g., PCT International Publication WO90/11364, published October 4, 1990

113. Sarver et al. (1990), *Science* 247:1222-1225

114. U.S. Patent No. 5,093,246

115. Haseloff and Gerlach (1988), *Nature* 334:585-591

116. Zaug et al. (1984), *Science* 224:574-578

117. Zaug and Cech (1986), *Science* 231:470-475

118. Zaug et al. (1986), *Nature* 324:429-433

119. Published International Application No. WO88/04300

120. Been and Cech (1986), *Cell* 47:207-216

121. Smithies et al. (1985), *Nature* 317:230-234

122. Thomas & Capecchi (1987), *Cell* 51:503-512

123. Thompson et al. (1989), *Cell* 5:313-321

124. Helene (1991), *Anticancer Drug Des.* 6:569-584

125. Helene et al. (1992), *Ann. N.Y. Acad. Sci.* 660:27-36

126. Maher. (1992), *Bioassays* 14:807-815

127. Goeddel (1990), Gene Expression Technology: Methods in Enzymology, *Academic Press*, San Diego, CA, p. 185

128. Broach et al. (1983) Experimental Manipulation of Gene Expression, M. Inouye, ed., *Academic Press*, p. 83

129. Molecular Cloning A Laboratory Manual (1989), 2nd Ed., ed. by Sambrook, Fritsch and Maniatis, eds., *Cold Spring Harbor Laboratory Press*, Chapters 16 and 17

130. Ben-Bassat et al. (1987) *J. Bacteriol.* 169:751-757

131. Miller et al. (1987), *Proc. Natl. Acad. Sci. USA* 84:2718-1722

132. EP Publication No: 0259149

133. Evans et al. (1989) *Nature* 339:385

134. Huang et al. (1988), *J. Virol.* 62:3855

135. Schlienger et al. (1992), *J. Virol.* 66:2

136. Posnett et al. (1988), *J. Biol. Chem.* 263:1719

137. Nardelli et al. (1992), *J. Immunol.* 148:914

138. See, e.g., Biochemistry, 2nd ed. (1981), Stryer, ed., *WH Freeman and Co.*

139. See, e.g., Narang (1983) *Tetrahedron* 39:3

140. Itakura et al. (1981) *Recombinant DNA, Proc. 3rd Cleveland Symposium*

141. See, e.g., *Macromolecules*, AG Walton, ed., *Elsevier Amsterdam* pp 273-289

142. Itakura et al. (1984), *Annu. Rev. Biochem.* 53:323

143. Itakura et al. (1984), *Science* 198:1056

144. Ike et al. (1983), *Nucl. Acid Res.* 11:477

145. Scott et al. (1990), *Science* 249:386-390

146. Roberts et al. (1992), *Proc. Natl. Acad. Sci., U.S.A.,*. 89:2429-2433

147. Devlin et al. (1990), *Science* 249: 404-406

148. Cwirla et al. (1990), *Proc. Natl. Acad. Sci.,U.S.A.,* 87: 6378-6382

149. U.S. Patents No. 5,223,409

150. U.S. Patents No. 5,198,346

151. U.S. Patents No. 5,096,815

152. Arkin and Yourvan (1992), *Proc. Natl. Acad. Sci. USA* 89:7811-7815

153. Yourvan et al. (1992), Parallel Problem Solving from *Nature*, 2

154. In Maenner and Manderick, eds., *Elsevier Publishing Co.*, Amsterdam, pp. 401-410

155. Delgrave et al. (1993), *Prot. Engineer.* 6:327-331

156. See Freidinger et al. (1988) in Peptides: Chemistry and Biology, G.R. Marshall, ed., *ESCOM,* Leiden, Netherlands

157. See Huffman et al. (1988) in Peptides: Chemistry and Biology, G.R. Marshall, ed., *ESCOM*, Leiden, Netherlands

158. Garvey et al. (1988) in Peptides: Chemistry and Biology, G.R. Marshall, ed., *ESCOM*, Leiden, Netherlands

159. Ewenson et al. (1986), *J. Med. Chem.* 29:295

160. Ewenson et al. in Peptides: Structure and Function, (Proceedings of the 9th American Peptide Symposium, *Pierce Chemical Co.* Rockland, IL, 1985)

161. Nagai et al. (1985), *Tetrahedron Lett.* 26:647

162. Sato et al. (1986), *J. Chem. Soc. Perkin. Trans.* 1:1231

163. Gordon et al. (1985), *Biochem. Biophys. Res. Commun.* 126:419

164. Dann et al. (1986), *Biochem. Biophys. Res. Commun.* 134:71

165. See, e.g., Antibodies: A Laboratory Manual, ed. by Harlow and Lane, *Cold Spring Harbor Press*, 1988

166. Kohler and Milstein (1975), *Nature 256*: 495-497

167. Kozbar et al. (1983), *Immunol. Today* 4: 72

168. Cole et al. (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96

169. Doetschman et al. (1985), *J. Embryol. Exp. Mol. Biol.* 87:27-45

170. *American Type Culture Collection*, Catalog no. CKL 1934

171. Ioffe et al. (1995), *Proc. Natl. Acad. Sci., USA* 92:7357-7361

172. Robertson in: Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, E.J. Robertson, ed., *IRL Press*, Washington, D.C., 1987

173. Bradley et al. (1986), *Current Topics Devel. Biol.* 20:357-371

174. Hogan et al. (Manipulating the Mouse Embryo: A Laboratory Manual (1986), *Cold Spring Harbor Laboratory Press*, Cold Spring Harbor, NY

175. Lakso et al. (1992), *Proc. Natl. Acad. Sci. USA* 89:6232-6236

176. Orban et al. (1992), *Proc. Natl. Acad. Sci. USA* 89:6861-6865

177. O'Gorman et al. (1991), *Science* 251:1351-1355

178. PCT publication WO 92/15694

179. Abremski et al. (1984), *J. Biol. Chem.* 259:1509-1514

180. Brinster et al. (1985), *Proc. Natl. Acad. Sci. USA* 82:4438-4442

181. Jaenich (1976), *Proc. Natl. Acad. Sci. USA* 73:1260-1264

182. Jahner et al. (1985), *Proc. Natl. Acad. Sci. USA* 82:6927-6931

183. Van der Putten et al. (1985), *Proc. Natl. Acad. Sci. USA* 82:6148-6152

184. Stewart et al. (1987), *EMBO J.* 6:383-388

185. Jahner et al. (1982), *Nature* 298:623-628

186. Evans et al. (1981), *Nature* 292:154-156

187. Bradley et al. (1984), *Nature* 309:255-258

188. Gossler et al. (1986), *Proc. Natl. Acad. Sci. USA* 83: 9065-9069

189. Robertson et al. (1986), *Nature* 322:445-448

190. Jaenisch (1988), *Science* 240:1468-1474

191. Habig et al (1974), *J. Biol. Chem.* 249:7130

192. Holmquist et al. (1979), *Anal. Biochem.* 95:540

193. U.S. Patent No. 5,259,045

194. Cronin et al. (1996), *Human Mutation* 7:244

195. See U.S. Patent No. 4,683,195

196. See U.S. Patent No. 4,683,202

197. See, e.g., Landegran et al. (1988), *Science* 241:1077-1080

198. See, e.g., Nakazawa et al. (1994), *Proc. Natl. Acad. Sci. USA* 91:360-364

199. See Abravaya et al. (1995), *Nuc. Acid. Res.* 23:675-682

200. Guatelli et al. (1990), *Proc. Natl. Acad. Sci. USA* 87:1874-1878

201. Kwoh et al. (1989), *Proc. Natl. Acad. Sci. USA* 86:1173-1177

202. Lizardi et al. (1988), *Bio/Technology* 6:1197

203. See, e.g., U.S. Patent No. 5,498,531

204. Maxim and Gilbert (1977), *Proc. Natl. Acad. Sci. USA,* 74:560

205. Sanger et al (1977), *Proc. Nat. Acad. Sci. USA,* 74:5463

206. *Biotechniques* (1995) 19:448

207. See, e.g., PCT publication WO 94/16101

208. Cohen et al. (1996), *Adv. Chromatogr.* 36:127-162

209. Griffin et al. (1993), *Appl. Biochem. Biotechnol.* 38:147-159

210. Myers, et al. (1985), *Science* 230:1242

211. See, e.g., Cotton et al (1988), *Proc. Natl. Acad. Sci. USA* 85:4397

212. Saleeba et al. (1992), *Meth. Enzymol.* 217:286-295

213. Hsu et al. (1994), *Carcinogenesis* 15:1657-1662

214. See, e.g., U.S. Patent No. 5,459,039

215. Orita et al. (1989), *Proc. Natl. Acad. Sci. USA* 86:2766

216. Cotton (1993), *Mutat. Res.* 285:125-144

217. Hayashi (1992), *Genet. Anal. Tech. Appl.* 9:73-79

218. Keen et al. (1991), *Trends Genet.* 7:5

219. Myers et al. (1985), *Nature* 313:495

220. Rosenbaum and Reissner (1987), *Biophys. Chem.* 265:12753

221. Saiki et al. (1986), *Nature* 324:163

222. Saiki et al (1989), *Proc. Natl. Acad. Sci. USA* 86:6230

223. Gibbs et al. (1989), *Nucl. Acids Res.* 17:2437-2448

224. Prossner (1993), *Tibtech* 11:238

225. Gasparini et al. (1992), *Mol. Cell Probes* 6:1

226. Barany (1991), *Proc. Natl. Acad. Sci. USA* 88:189

227. See U.S. Pat. No. 4,998,617

228. Nickerson et al. (1990), *Proc. Natl. Acad. Sci. USA* 87:8923-8927

229. Tobe et al. (1996), *Nucleic Acids Res.* 24: 3728

230. U.S. Pat. No.4,656,127

231. French Patent 2,650,840; PCT Appln. No. WO91/02087

232. PCT Appln. No. 92/15712

233. Komher et al. (1989), *Nucl. Acids. Res.* 17:7779-7784

234. Sokolov (1990), *Nucl. Acids Res.* 18:3671

235. Syvanen, et al. (1990), *Genomics* 8:684-692

236. Kuppuswamy et al. (1991), *Proc. Natl. Acad. Sci. USA* 88:1143-1147

237. Prezant et al. (1992), *Hum. Mutat.* 1:159-164

238. Ugozzoli et al. (1992), *GATA* 9:107-112

239. Nyren et al. (1993), *Anal. Biochem.* 208:171-175

240. Syvanen et al. (1993), *Amer. J. Hum. Genet.* 52:46-59

241. Roest, et. al. (1993), *Hum. Mol. Genet.* 2:1719-21

242. van der Luijt, et. al. (1994), *Genomics* 20:1-4

243. Voller (1978), *Diagnostic Horizons* 2:1-7

244. Voller et al. (1978), *J. Clin. Pathol.* 31:507-520

245. Butler (1981), *Meth. Enzymol.* 73:482-523

246. Maggio (ed.) (1980), Enzyme Immunoassay, *CRC Press*, Boca Raton, FL

247. Ishikawa, et al. (eds.) (1981), Enzyme Immunoassay, *Kgaku Shoin*, Tokyo

248. See, e.g., Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, *The Endocrine Society*, March, 1986

249. LaVail et al. (1998), *Invest. Ophthalmol. Vis. Sci.* 39:592-602

250. La Vail et al. (1998), *Invest. Ophthalmol. Vis. Sci.* 39:581-591

251. See U.S. Patent 5,328,470

252. Chen et al. (1994), *Proc. Natl. Acad. Sci., USA* 91: 3054-3057

253. Dev et al. (1994), *Cancer Treat. Rev.* 20:105-115

254. *DNA Cloning*, Volumes I and II, Glover, ed., 1985

255. Oligonucleotide Synthesis, M. J. Gait, ed., 1984

256. Nucleic Acid Hybridization, D. Hames & S. J. Higgins, eds., 1984

257. Transcription and Translation, B. D. Hames & S. J. Higgins, eds., 1984

258. Culture Of Animal Cells, R. I. Freshney, *Alan R. Liss, Inc.*, 1987

259. Immobilized Cells And Enzymes, IRL Press, 1986

260. Perbal (1984), A Practical Guide To Molecular Cloning

261. See Methods In Enzymology (*Academic Press, Inc.*, N.Y.)

262. Gene Transfer Vectors For Mammalian Cells, J. H. Miller and M. P. Calos, eds., *Cold Spring Harbor Laboratory*, 1987

263. Methods In Enzymology, Vols. 154 and 155, Wu et al., eds., *Academic Press Inc.*, N.Y.

264. Immunochemical Methods In Cell And Molecular Biology (Mayer and Walker, eds., *Academic Press*, London, 1987

265. Handbook Of Experimental Immunology, Volumes I-IV, D. M. Weir and C. C. Blackwell, eds., 1986

266. Williams and Barclay (1988), *Ann. Rev. Immun.* 6:381-405

267. Krusius et al. (1987), *J. Biol. Chem.* 262:13120-13125

268. Sanger et al. (1977), *Proc. Natl. Acad. Sci. USA* 74:5463-5467

269. Larrick (1992), *Trends Biotechnol.* 10:146-152

270. Wang et al. (1989), *Proc. Natl. Acad. Sci. U.S.A.* 86:9717-9721

271. Staecker et al. (1994), *BioTechniques* 16:76-80

272. Golde et al. (1990), *Neuron* 4:253-267

273. Chirgwin et al. (1979), *Biochemistry* 18:5294-5299

274. Chomczynski and Sacchi (1987), *Anal. Biochem.* 162:156-159

275. Akiyama et al. (1989), *J. Cell Biol.* 109:863-875

276. Hayashi et al. (1992), *J. Cell Biology* 119:945-959

277. Ruoslahti (1988), *Annual Rev. Cell Biol.* 4:229-255

278. Wayner et al. (1991), *J. Cell Biol.* 113:919-929